



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 129485

**TO: Mary Mosher
Location: REM/3D25/3C18
Art Unit : 1648
Monday, August 23, 2004

Case Serial Number: 09/851410**

**From: Toby Port
Location: Biotech-Chem Library
REM-1A59
Phone: (571) 272-2523
toby.port@uspto.gov**

Search Notes

Blank sheet

STIC-Biotech/ChemLib

159485

From: Mosher, Mary
Sent: Tuesday, August 10, 2004 8:17 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search 09/851410

Please translate & search AA encoded by:

nt 5147-7129 of SEQ 6

nt 5117-7096 of SEQ 10

And search SEQ 8.

Please do regular search, oligo search for the seq 6 & 10 translation products, and interference search. Thanks.

Mary Mosher
AU 1648
571-272-0906
Office Rem 3D25
Mail Rem 3C18

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

Blank sheet

source	1. .1983			
	/organism="unknown"			
	/mol_type="unassigned DNA"			
ORIGIN				
Query Match	100.0%;	Score 1983;	DB 6;	Length 1983;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1983;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCARGTTTTTGGCCCTATGCTCCCGCGCCA	60	
Db	1	ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTTGGCCCTATGCTCCCGCGCCA	60	
QY	61	CGCGCCGGTCAAGCTCTGCGCGCGCGTCTGTGGGCGGCGCAGCGCGGTGGT	120	
Db	61	CGCGCCGGTCAAGCTCTGCGCGCGCGTCTGTGGGCGGCGCAGCGCGGTGGT	120	
QY	121	TTCTGGGGTGACCGGGTTGATTTCTCAGCGCTTCGCAATCCCTATATTCATCCAAACAC	180	
Db	121	TTCTGGGGTGACCGGGTTGATTTCTCAGCGCTTTGCAATCCCTCTATATTCATCCAAACAC	180	
QY	181	CCCTTCGCCCGCGATGTCAACGCTGCGCGCGGGGCTGGACCTCGTGTTCGCCAACCCGCC	240	
Db	181	CCCTTCGCCCGCGATGTCAACGCTGCGCGCGGGGCTGGACCTCGTGTTCGCCAACCCGCC	240	
QY	241	CGACCACTCGGCTCGGCTTGGGTGACCAAGGCCAGCGCCCGCGGTGCTCAAGTCGT	300	
Db	241	CGACCACTCGGCTCGGCTTGGGTGACCAAGGCCAGCGCCCGCGGTGCTCAAGTCGT	300	
QY	301	AGACCTTACCAAGCTGGGGCGCGCGCTTAACCGGGTTCGCTCGCGGCCCATGACACCCCG	360	
Db	301	AGACCTTACCAAGCTGGGGCGCGCGCTTAACCGGGTTCGCTCGCGGCCCATGACACCCCG	360	
QY	361	CCAGTGCCTGATGTCGACTCCCGCGCGCCATCTTTGGCGCGGCAATACCTATCAACA	420	
Db	361	CCAGTGCCTGATGTCGACTCCCGCGCGCCATCTTTGGCGCGGCAATACCTATCAACA	420	
QY	421	TCTCCCTTACCTCTTCCGTGGGCACCGGCACTAACCTGGTTCTTTATGCGGCCCTCTT	480	
Db	421	TCTCCCTTACCTCTTCCGTGGGCACCGGCACTAACCTGGTTCTTTATGCGGCCCTCTT	480	
QY	481	AGTCGGCTTTTACCCCTTCAGGACGGCACCAATACCATATATGGCCACGGAGCTTCT	540	
Db	481	AGTCGGCTTTTACCCCTTCAGGACGGCACCAATACCATATATGGCCACGGAGCTTCT	540	
QY	541	AATTATGCCAGTACCGGGTTGCCGTGCGACCAATCGTTACCGCCCGCTGCTCCCAAT	600	
Db	541	AATTATGCCAGTACCGGGTTGCCGTGCGACCAATCGTTACCGCCCGCTGCTCCCAAT	600	
QY	601	GCTGTGGCGGTTACGGCATCTCCATCTCATPTCTGGCGACAGACCAACCCCGAAG	660	
Db	601	GCTGTGGCGGTTACGGCATCTCCATCTCATPTCTGGCGCACAGACCAACCCCGGACG	660	
QY	661	TCCGTTGATATGAATTCATAAACCCTCGACGATGTTCTGATTTTATGTCAGCGCGGATA	720	
Db	661	TCCGTTGATATGAATTCATAAACCCTCGACGATGTTCTGATTTTATGTCAGCGCGGATA	720	
QY	721	GCCTCTGAGCTTGTGATCCCAAGTACGCGCTACACTATCGTAACCAAGGCTGGCGCTCC	780	
Db	721	GCCTCTGAGCTTGTGATCCCAAGTACGCGCTACACTATCGTAACCAAGGCTGGCGCTCC	780	
QY	781	GTCCAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCTGGTCTGTTATGCTTTGCATA	840	
Db	781	GTCCAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTGTTATGCTTTGCATA	840	
QY	841	CATGGCTCACTCGTAAATTCCTATCTAATAACACCTTATACCGGTGCCCTCGGGCTGTTG	900	
Db	841	CATGGCTCACTCGTAAATTCCTATCTAATAACACCTTATACCGGTGCCCTCGGGCTGTTG	900	
QY	901	GACTTTGCCCTTGAGCTTGATTTTCGCAACCTTACCCCGGTTAAACCAATACCGGGTTC	960	
Db	901	GACTTTGCCCTTGAGCTTGATTTTCGCAACCTTACCCCGGTTAAACCAATACCGGGTTC	960	

QY	961	TCCCGTTATTCCAGCACTGCTCGCCACCGCCTTGGTCGCGGTGCGGACGGGACTGCCGAG	1020
Db	961	TCCCGTTATTCCAGCACTGCTCGCCACCGCCTTGGTCGCGGTGCGGACGGGACTGCCGAG	1020
QY	1021	CTCACCACACGGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT	1080
Db	1021	CTCACCACACGGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT	1080
QY	1081	GTCCGGTGAGATCGGCGCGGGATAGCCCTCACCCCTGTTCAACCTTGTCTGACACTCTGCTT	1140
Db	1081	GTCCGGTGAGATCGGCGCGGGATAGCCCTCACCCCTGTTCAACCTTGTCTGACACTCTGCTT	1140
QY	1141	GGCGGCTCCCGACAGAATTGATTTGTCGGCTGGTGGCCAGCTGTTCTACTCCCGTCCC	1200
Db	1141	GGCGGCTCCCGACAGAATTGATTTGTCGGCTGGTGGCCAGCTGTTCTACTCCCGTCCC	1200
QY	1201	GTGTCTCAGCCAAATGGCGAGCCGACTGTTAAAGTTGTATACATCTGTGTAGAGAACTCTCAG	1260
Db	1201	GTGTCTCAGCCAAATGGCGAGCCGACTGTTAAAGTTGTATACATCTGTGTAGAGAACTCTCAG	1260
QY	1261	CAGGATAAAGGTAATTGCAATCCGCATGACATTTGACCTCGGAGAAATCTCTGTGTGTTATT	1320
Db	1261	CAGGATAAAGGTAATTGCAATCCGCATGACATTTGACCTCGGAGAAATCTCTGTGTGTTATT	1320
QY	1321	CAGGATTATGATAACCAACATGAACAGATCGGCCAGCCCTTCTCCAGCCCCCATCGCGC	1380
Db	1321	CAGGATTATGATAACCAACATGAACAGATCGGCCAGCCCTTCTCCAGCCCCCATCGCGC	1380
QY	1381	CTTTTCTCTGTCCTTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCCAGTAT	1440
Db	1381	CTTTTCTCTGTCCTTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCCAGTAT	1440
QY	1441	GACCACTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG	1500
Db	1441	GACCACTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG	1500
QY	1501	GTTAATGTTTGGACCGCGCGCAGGCGTTCGCCGGTTCGTCGATTTGGACCAAGGTACACA	1560
Db	1501	GTTAATGTTTGGACCGCGCGCAGGCGTTCGCCGGTTCGTCGATTTGGACCAAGGTACACA	1560
QY	1561	CTTGAGGTTCCGCCCTCTCCACCATTCAGCAGTACTCGAAGACGTTCTTTTGTCTGCCG	1620
Db	1561	CTTGAGGTTCCGCCCTCTCCACCATTCAGCAGTACTCGAAGACGTTCTTTTGTCTGCCG	1620
QY	1621	CTCCCGGTAAGCTCTCTTTCTGGGAGGACGACCACTAAAGCGGGTACCTTTATTAAT	1680
Db	1621	CTCCCGGTAAGCTCTCTTTCTGGGAGGACGACCACTAAAGCGGGTACCTTTATTAAT	1680
QY	1681	TATAACACACTGTAGCGACCAACTGCTGTGAGAAATGCGCGGGSCACCGGGTCTGCT	1740
Db	1681	TATAACACACTGTAGCGACCAACTGCTGTGAGAAATGCGCGGGSCACCGGGTCTGCT	1740
QY	1741	ATTTCACACTTACCACTAGCTGGGTGCTGGTCCGCTCTCAATTTCTGCGGTTGCCGTT	1800
Db	1741	ATTTCACACTTACCACTAGCTGGGTGCTGGTCCGCTCTCAATTTCTGCGGTTGCCGTT	1800
QY	1801	TTAGCCCCCCTCTGCGCTAGCATTTGCTTGAAGATACCTTGGACTACCTGCCCCGCGC	1860
Db	1801	TTAGCCCCCCTCTGCGCTAGCATTTGCTTGAAGATACCTTGGACTACCTGCCCCGCGC	1860
QY	1861	CATACCTTTGATGATTTCTGCCAGAGTGCGCCCTTGGCCTTCAGGGCTGCGCTTC	1920
Db	1861	CATACCTTTGATGATTTCTGCCAGAGTGCGCCCTTGGCCTTCAGGGCTGCGCTTC	1920
QY	1921	CAGTCTACTGTGCTGAGCTTCAGCGCTTAAGATGAGGTGGGTAAACCTCGGGAGTTG	1980
Db	1921	CAGTCTACTGTGCTGAGCTTCAGCGCTTAAGATGAGGTGGGTAAACCTCGGGAGTTG	1980
QY	1981	TAG 1983	
Db	1981	TAG 1983	

RESULT 2
AR002160
LOCUS AR002160 2049 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 3 from patent US 5741490.
ACCESSION AR002160
VERSION AR002160.1 GI:3963714
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
UNCLASSIFIED.
REFERENCE 1 (bases 1 to 2049)
AUTHORS Reyes, G.R., Bradley, D.W., Twu, J.-S., Purdy, M.A., Tam, A.W.,
Krawczynski, K.Z. and Yarbrough, P.D.
TITLE Hepatitis B virus vaccine and method
JOURNAL Patent: US 5741490-A 3 21-APR-1998;
FEATURES
Location/Qualifiers
1..2049
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 1983; DB 6; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGCTCGGCTATTTTGTGCTGCTCCATGCTTTTGCCTATGCTGCCCGGCCA 60
Db 1 ATGGGCGCTCGGCTATTTTGTGCTGCTCCATGCTTTTGCCTATGCTGCCCGGCCA 60
QY 61 CCGCGCGGTGACGGGTCTCGGCGCGCTCGTGGCGGCGGAGCGGCTTCCGCGGTGGT 120
Db 61 CCGCGCGGTGACGGGTCTCGGCGCGCTCGTGGCGGCGGAGCGGCTTCCGCGGTGGT 120
QY 121 TTCCTGGGTGACGGGTGTAATCTAGCCCTTCGCAATCCCTATATTCATCCAAAC 180
Db 121 TTCCTGGGTGACGGGTGTAATCTAGCCCTTCGCAATCCCTATATTCATCCAAAC 180
QY 181 CCCTTCGCGCGCGATGTCACCGCTCGCGCGCGGCTGAGCTCGTGTGCGCAACCGCC 240
Db 181 CCCTTCGCGCGCGATGTCACCGCTCGCGCGCGGCTGAGCTCGTGTGCGCAACCGCC 240
QY 241 CGACCACTCGGCTCCGCTTGGGTGACACGCGCGCGGCTGAGCTCGTGTGCGCAACCGCC 300
Db 241 CGACCACTCGGCTCCGCTTGGGTGACACGCGCGCGGCTGAGCTCGTGTGCGCAACCGCC 300
QY 301 AGACCTACACAGCTGGGCGCGCGCTAAACCGCGGTGCTCGCGCGCGGCTGAGCTCGTGT 360
Db 301 AGACCTACACAGCTGGGCGCGCGCTAAACCGCGGTGCTCGCGCGCGGCTGAGCTCGTGT 360
QY 361 CCAGTGCCCTGATGTCAGCTCCGCGCGCGGCTAAACCGCGGTGCTCGCGCGCGGCTGAGCT 420
Db 361 CCAGTGCCCTGATGTCAGCTCCGCGCGCGGCTAAACCGCGGTGCTCGCGCGCGGCTGAGCT 420
QY 421 TCTCCGCTTACCTCTTCGCTGGGCGACGCGGCTAAACCGCGGTGCTCGCGCGCGGCTGAGCT 480
Db 421 TCTCCGCTTACCTCTTCGCTGGGCGACGCGGCTAAACCGCGGTGCTCGCGCGCGGCTGAGCT 480
QY 481 AGTCCGCTTTACCTCTTCGCTGGGCGACGCGGCTAAACCGCGGTGCTCGCGCGCGGCTGAGCT 540
Db 481 AGTCCGCTTTACCTCTTCGCTGGGCGACGCGGCTAAACCGCGGTGCTCGCGCGCGGCTGAGCT 540
QY 541 AATTATGCCAGTACCGGTTGCCGTTGCCGTTGCCGTTGCCGTTGCCGTTGCCGTTGCCGTT 600
Db 541 AATTATGCCAGTACCGGTTGCCGTTGCCGTTGCCGTTGCCGTTGCCGTTGCCGTTGCCGTT 600
QY 601 GCTGTCCGCGGTTACCGCTCTCCATCTCATCTGCGCGACAGACCAACCAACCGCGGAGC 660
Db 601 GCTGTCCGCGGTTACCGCTCTCCATCTCATCTGCGCGACAGACCAACCAACCGCGGAGC 660
QY 661 TCCGTTGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
Db 661 TCCGTTGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720

QY 721 GCTCTGAGCTTGTGATCCAAAGTGAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
Db 721 GCTCTGAGCTTGTGATCCAAAGTGAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
Db 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
QY 841 CATGGCTACTCGTAATTCCTATACTATAACCTATACCGGTGCGCTGGGCTGTG 900
Db 841 CATGGCTACTCGTAATTCCTATACTATAACCTATACCGGTGCGCTGGGCTGTG 900
QY 901 GACTTTGCCCTTGAGCTTGAGTTTGCACACTTACCCCGGTAAACCAATACCGCGGTC 960
Db 901 GACTTTGCCCTTGAGCTTGAGTTTGCACACTTACCCCGGTAAACCAATACCGCGGTC 960
QY 961 TCCGCTTATTCAGCACTGCTCGCCACCGCTTCTGTCGGGTGCGGACGGGACTGCCAG 1020
Db 961 TCCGCTTATTCAGCACTGCTCGCCACCGCTTCTGTCGGGTGCGGACGGGACTGCCAG 1020
QY 1021 CTCACCAACCGCTGCTACCGCTTATGAGGACCTCTATTTACTAGTACTAATGCT 1080
Db 1021 CTCACCAACCGCTGCTACCGCTTATGAGGACCTCTATTTACTAGTACTAATGCT 1080
QY 1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCCCTGTTCAACCTTCTGCTGACACTCTGCT 1140
Db 1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCCCTGTTCAACCTTCTGCTGACACTCTGCT 1140
QY 1141 GCGCGCTTCCGCAAGAAATTTGTCGGGTGGTGGCGAGCTGTTCTACTCCCGTCC 1200
Db 1141 GCGCGCTTCCGCAAGAAATTTGTCGGGTGGTGGCGAGCTGTTCTACTCCCGTCC 1200
QY 1201 GTTGTCTAGCAATGCGCGAGCGACTGTAAAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db 1201 GTTGTCTAGCAATGCGCGAGCGACTGTAAAGTTGTATACATCTGTAGAGATGCTCAG 1260
QY 1261 CAGGATAGGGTATTTGCAATCCCGATGACATTTGACCTCGGAGAACTCTCGTGTGTTAT 1320
Db 1261 CAGGATAGGGTATTTGCAATCCCGATGACATTTGACCTCGGAGAACTCTCGTGTGTTAT 1320
QY 1321 CAGGATATGATAACCAACATGAAGATCGGCGAGCGCTTCTCCAGCCCAATCGCG 1380
Db 1321 CAGGATATGATAACCAACATGAAGATCGGCGAGCGCTTCTCCAGCCCAATCGCG 1380
QY 1381 CTTTCTCTGCTGCTGAGCTAATGATGCTTGGCTCTCTCTACCGCTGCGAGTAT 1440
Db 1381 CTTTCTCTGCTGCTGAGCTAATGATGCTTGGCTCTCTCTACCGCTGCGAGTAT 1440
QY 1441 GACCACTCCACTTATGGCTCTTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACTTG 1500
Db 1441 GACCACTCCACTTATGGCTCTTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACTTG 1500
QY 1501 GTTAATGTTGCAACCGCGCGAGCGCTTGGCGCGTCTCGATTTGGAACCAAGTCA 1560
Db 1501 GTTAATGTTGCAACCGCGCGAGCGCTTGGCGCGTCTCGATTTGGAACCAAGTCA 1560
QY 1561 CTTGAGGTTGCGCGCTCTCCAGCACTCCAGAGTACTCGAAGACCTTCTTGTCTGCGG 1620
Db 1561 CTTGAGGTTGCGCGCTCTCCAGCACTCCAGAGTACTCGAAGACCTTCTTGTCTGCGG 1620
QY 1621 CTCGCGGTAAAGCTCTCTTCTGGGAGGAGGAGCAAACTAAAGCGGGTACCTTATAAT 1680
Db 1621 CTCGCGGTAAAGCTCTCTTCTGGGAGGAGGAGCAAACTAAAGCGGGTACCTTATAAT 1680
QY 1681 TATAACCACTGCTAGCGAACCACTGCTTGTGAGAAATGCCCGCGGACCGGGTCT 1740
Db 1681 TATAACCACTGCTAGCGAACCACTGCTTGTGAGAAATGCCCGCGGACCGGGTCT 1740
QY 1741 ATTTCCACTTACCACTAGCGTGGTGGTCTCGCTCCATTTCTGCGGTGCGGTT 1800
Db 1741 ATTTCCACTTACCACTAGCGTGGTGGTCTCGCTCCATTTCTGCGGTGCGGTT 1800
QY 1801 TTAGCCCCCACTCTGCGCTAGCAATTTGTTGAGGATACCTTGGACTACCTTCCCGCGC 1860

```
Db 1801 TTAGCCCCCACTCTGCGTAGCATTTGCTTGAGATACCTTGGACTACCTGCGCGCC 1860
Qy 1861 CAFACTTTTGATGATTTTCGCCAGAGTGCGCCCTTGCGCTTCAGGGGTGCGCTTTC 1920
Db 1861 CATACTTTTGATGATTTTCGCCAGAGTGCGCCCTTGCGCTTCAGGGGTGCGCTTTC 1920
Qy 1921 CAGTCTACTGCGCTGAGCTTACGCGCTTAAGATGAAGTGGGTAAACTCGGAGTTG 1980
Db 1921 CAGTCTACTGCGCTGAGCTTACGCGCTTAAGATGAAGTGGGTAAACTCGGAGTTG 1980
Qy 1981 TAG 1983
Db 1981 TAG 1983

RESULT 3
ARI170404
LOCUS ARI170404 2049 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 1 from patent US 6291641.
ACCESSION ARI170404
VERSION ARI170404.1 GI:17908363
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 2049)
AUTHORS Fuerst, R.K., McAtree, C., Patrick, Y., Yarbrough, P.O. and Zhang, Y.-F.
TITLE Hepatitis E virus antigens and uses therefor
JOURNAL Patent: US 6291641-A 1 18-SEP-2001;
FEATURES
source
1. 2049
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 1983; DB 6; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTTGTGCTATGCTGCGCGCCA 60
Db 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTTGTGCTATGCTGCGCGCCA 60
Qy 61 CCGCCCGGTGAGCCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGT 120
Db 61 CCGCCCGGTGAGCCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGT 120
Qy 121 TTCTGGGGTGACCGGGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 180
Db 121 TTCTGGGGTGACCGGGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 180
Qy 181 CCTTTCGCCCGCGATGTCACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 CCTTTCGCCCGCGATGTCACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Qy 241 CGACCACTCGCTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 300
Db 241 CGACCACTCGCTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 300
Qy 301 AGACCTACCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 AGACCTACCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy 361 CCAGTGCTGATGTCAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Db 361 CCAGTGCTGATGTCAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Qy 421 TCTCCCTTACTCTTCTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 421 TCTCCCTTACTCTTCTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Qy 481 AGTCCGCTTTTACCCCTTCAGGACGGCCCAATACCCATATAATGGCCACGGAAGCTTCT 540
```

```
Db 481 AGTCCGCTTTTACCCCTTCAGGACGGCCCAATACCCATATAATGGCCACGGAAGCTTCT 540
Qy 541 AATATATCCAGATACCGGGTGGCGGTGCACAAATCCGTTACGCCCGCTGGTCCCAAT 600
Db 541 AATATATCCAGATACCGGGTGGCGGTGCACAAATCCGTTACGCCCGCTGGTCCCAAT 600
Qy 601 GCTGTGCGCGGTACGCCATCTCCATCTCTGCGCCACAGACCAACACCCCGCGCG 660
Db 601 GCTGTGCGCGGTACGCCATCTCCATCTCTCTGCGCCACAGACCAACACCCCGCGCG 660
Qy 661 TCCGTTGATATGAATTCATAAACCCTCGACGGATGTTGTTATTTAGTCCAGCCCGGATA 720
Db 661 TCCGTTGATATGAATTCATAAACCCTCGACGGATGTTGTTATTTAGTCCAGCCCGGATA 720
Qy 721 GCCTCTGAGCTTGTGATCCCAAGTGGCGCTACACTATCGTAAACAGGCTGGCGTCC 780
Db 721 GCCTCTGAGCTTGTGATCCCAAGTGGCGCTACACTATCGTAAACAGGCTGGCGTCC 780
Qy 781 GTGAGACCTCTGCGGTGGCTGAGGAGGAGGCTACCTCTGCTCTGTTATGCTTTGCATA 840
Db 781 GTGAGACCTCTGCGGTGGCTGAGGAGGAGGCTACCTCTGCTCTGTTATGCTTTGCATA 840
Qy 841 CATGGCTCACTCGTAAATTCCTATACCTATAACCTTATACCGTGGCTGGCGCTGTG 900
Db 841 CATGGCTCACTCGTAAATTCCTATACCTATAACCTTATACCGTGGCTGGCGCTGTG 900
Qy 901 GACTTTCCTTGGCTGAGTTTGGCAACTTACCCCGGTAAACCAATACCGCGGTC 960
Db 901 GACTTTCCTTGGCTGAGTTTGGCAACTTACCCCGGTAAACCAATACCGCGGTC 960
Qy 961 TCCGTTTATTCAGACCTGCTCGCCACCGCTTCTGCTCGCGTGGCGAGGACTGCGCG 1020
Db 961 TCCGTTTATTCAGACCTGCTCGCCACCGCTTCTGCTCGCGTGGCGAGGACTGCGCG 1020
Qy 1021 CTCACCAACCGCTGCTACCGCTTATGAAGACCTCTATTTTACTAGTAAATGGT 1080
Db 1021 CTCACCAACCGCTGCTACCGCTTATGAAGACCTCTATTTTACTAGTAAATGGT 1080
Qy 1081 GTGCGTGAATCGCGCGGATAGCCCTCACCTGTTTCAACCTTGTGACACTGCTT 1140
Db 1081 GTGCGTGAATCGCGCGGATAGCCCTCACCTGTTTCAACCTTGTGACACTGCTT 1140
Qy 1141 GCGCGCTGCGCGAGAAATGATTTCGTCGCTGGTGGCGAGCTGTTCTACTCCCGTCCC 1200
Db 1141 GCGCGCTGCGCGAGAAATGATTTCGTCGCTGGTGGCGAGCTGTTCTACTCCCGTCCC 1200
Qy 1201 GTTGTCTCAGCAATGCGAGCGACTGTTAAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
Db 1201 GTTGTCTCAGCAATGCGAGCGACTGTTAAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
Qy 1261 CAGGATGAAGGTTTGAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT 1320
Db 1261 CAGGATGAAGGTTTGAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT 1320
Qy 1321 CAGGATTTATGAATCAACCAATGAACATGACGCGCGCTTCTCCAGCCCGCATCGCGC 1380
Db 1321 CAGGATTTATGAATCAACCAATGAACATGACGCGCGCTTCTCCAGCCCGCATCGCGC 1380
Qy 1381 CCTTTCTCTGCTTTCAGCTAAATGATGCTTTGGCTCTCTCTCAACCGTGGCGAGTAT 1440
Db 1381 CCTTTCTCTGCTTTCAGCTAAATGATGCTTTGGCTCTCTCTCAACCGTGGCGAGTAT 1440
Qy 1441 GACCGTCCACTTATGCTCTGACGCGCGCGCTTGGCGCGCTGCTGCTGCTGCTGCTGCT 1500
Db 1441 GACCGTCCACTTATGCTCTGACGCGCGCGCTTGGCGCGCTGCTGCTGCTGCTGCTGCT 1500
Qy 1501 GTTAAATGTCGACCGCGCGCGCGCTTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db 1501 GTTAAATGTCGACCGCGCGCGCGCTTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Qy 1561 CTTGACCGCTGCGCGCGCTTCTCAACATCCAGAGTACTCGAAGACCTTCTTTGCTGCGCG 1620
```

```
Db 1561 CTTGACGGTGGCCCTCTCCACCATCCAGAGTACTCGAAGACCTCTTTTGCTCGCG 1620
Qy 1621 CTCGCGGTAGCTCTCTTTCTGGAGGAGGAGCAAACTAAAGCCGGGTACCTTTATAT 1680
Db 1621 CTCGCGGTAGCTCTCTTTCTGGAGGAGGAGCAAACTAAAGCCGGGTACCTTTATAT 1680
Qy 1681 TATAACACACGTAGAGGAGCAAACTGCTTGTGCGAGAAATGCCGCGGACCGGCTCGCT 1740
Db 1681 TATAACACACGTAGAGGAGCAAACTGCTTGTGCGAGAAATGCCGCGGACCGGCTCGCT 1740
Qy 1741 ATTTCCACTTACACACGTAGAGGAGCAAACTGCTTGTGCGAGAAATGCCGCGGACCGGCTCGCT 1800
Db 1741 ATTTCCACTTACACACGTAGAGGAGCAAACTGCTTGTGCGAGAAATGCCGCGGACCGGCTCGCT 1800
Qy 1801 TTAGCGCCCATCTGCGGTAGCATGCTTGGAGATACCTTGGACTACCTGCGCGGCG 1860
Db 1801 TTAGCGCCCATCTGCGGTAGCATGCTTGGAGATACCTTGGACTACCTGCGCGGCG 1860
Qy 1861 CATACCTTTGATGATTTCTGCCAGAGTGCAGCGCCCTTGGCGCTTCAGGGCTGCGCTTTC 1920
Db 1861 CATACCTTTGATGATTTCTGCCAGAGTGCAGCGCCCTTGGCGCTTCAGGGCTGCGCTTTC 1920
Qy 1921 CAGCTACTGCTGCTGAGCTTCCAGCGCTTAAAGATGAAGGTGGTAAACTCGGAGTTG 1980
Db 1921 CAGCTACTGCTGCTGAGCTTCCAGCGCTTAAAGATGAAGGTGGTAAACTCGGAGTTG 1980
Qy 1981 TAG 1983
Db 1981 TAG 1983

RESULT 4
AR232465
LOCUS AR232465 2049 bp mRNA linear PAT 20-DEC-2002
DEFINITION Sequence 3 from patent US 6455492.
ACCESSION AR232465
VERSION AR232465.1 GI:27274534
KEYWORDS
SOURCE Unknown.
ORIGIN
Unclassified.
REFERENCE 1 (bases 1 to 2049)
AUTHORS Reyes, G.R., Bradley, D.W., Twu, J.-S., Purdy, M.A., Tam, A.W.,
Krawczynski, K.Z. and Yarbough, P.O.
Hepatitis B virus vaccine and method
Patent: US 6455492-A 3 24-SEP-2002;
FEATURES
Location/Qualifiers
1..2049
/organism="unknown"
/mol_type="mRNA"
ORIGIN
Query Match 100.0%; Score 1983; DB 6; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCCTATGTCGCCGCGCA 60
Db 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCCTATGTCGCCGCGCA 60
Qy 61 CCGCCCGGTGACCGGCTGCTGCGCGCGCTGCTGCGCGCGAGCGGCGGTTCGCGCGTGT 120
Db 61 CCGCCCGGTGACCGGCTGCTGCGCGCGCTGCTGCGCGCGAGCGGCGGTTCGCGCGTGT 120
Qy 121 TTCTGGGTGACCGGCTGATTTCTCAGCCCTTCGCAATCCCTTATATTCATCAACCAAC 180
Db 121 TTCTGGGTGACCGGCTGATTTCTCAGCCCTTCGCAATCCCTTATATTCATCAACCAAC 180
Qy 181 CCTTCGCCCGCATGTCAGCGCTGCGCGCGGCGGTGAGCTGCTGTCGCAACCGCGC 240
Db 181 CCTTCGCCCGCATGTCAGCGCTGCGCGCGGCGGTGAGCTGCTGTCGCAACCGCGC 240
Qy 241 CGACCACTCGGCTCGCTTGGCGTGACAGGCCGCGCGCGCTTGCCTCAGCTCGT 300
```

```
Db 241 CGACCACTCGGCTCGCTTGGCGTGACAGGCCCGCGCGCTTGCCTCAGCTCGT 300
Qy 301 AGACCTACACAGCTGGGGCGCGCGCTAACCGCGGTGCTCCGCGCCATACACCCCG 360
Db 301 AGACCTACACAGCTGGGGCGCGCGCTAACCGCGGTGCTCCGCGCCATACACCCCG 360
Qy 361 CCAGTGCCTGATGTCGACTCCCGCGGCGCACTCTTTGGCGCGGAGTATAACCTATCAACA 420
Db 361 CCAGTGCCTGATGTCGACTCCCGCGGCGCACTCTTTGGCGCGGAGTATAACCTATCAACA 420
Qy 421 TCTCCCTCTTACCTCTTCCGTGCGCAACCGGCACTAACTGGTCTTTTATGCGCCCTCTT 480
Db 421 TCTCCCTCTTACCTCTTCCGTGCGCAACCGGCACTAACTGGTCTTTTATGCGCCCTCTT 480
Qy 481 AGTCGCTTTTACCTCTTCCGTGCGCAACCGGCACTAACTGGTCTTTTATGCGCCCTCTT 540
Db 481 AGTCGCTTTTACCTCTTCCGTGCGCAACCGGCACTAACTGGTCTTTTATGCGCCCTCTT 540
Qy 541 AATTATGCCAGTACCGGCTTCCCGCTGCCAATCCGTTACCGCGCTGGTCCCAAT 600
Db 541 AATTATGCCAGTACCGGCTTCCCGCTGCCAATCCGTTACCGCGCTGGTCCCAAT 600
Qy 601 GCTGTCGGCGGTACGCCATCTCCATCTCATTTCTGCGCAACAGACCAACACCCGCG 660
Db 601 GCTGTCGGCGGTACGCCATCTCCATCTCATTTCTGCGCAACAGACCAACACCCGCG 660
Qy 661 TCCGTTGATATGAATCAATTAACCTCGAGCGGATGTTTCTGTTATGTCAGCGCGCAT 720
Db 661 TCCGTTGATATGAATCAATTAACCTCGAGCGGATGTTTCTGTTATGTCAGCGCGCAT 720
Qy 721 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTTACACTATCGTAACCAAGCTGGCGTCC 780
Db 721 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTTACACTATCGTAACCAAGCTGGCGTCC 780
Qy 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTA CTTCTGCTGTTTGTATGCTTGCATA 840
Db 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTA CTTCTGCTGTTTGTATGCTTGCATA 840
Qy 841 CATGCTCACTCGTAAATTCCTATATACCTATACCGCTTACCGGTCCTCGGCTGTTG 900
Db 841 CATGCTCACTCGTAAATTCCTATATACCTATACCGGTCCTCGGCTGTTG 900
Qy 901 GACTTTGCCCTTGAAGCTTGAAGTTTCGCAACCTTACCGCGGTAAACACCAATACGCGGTC 960
Db 901 GACTTTGCCCTTGAAGCTTGAAGTTTCGCAACCTTACCGCGGTAAACACCAATACGCGGTC 960
Qy 961 TCCGTTATTCAGCACTGCTCGCCACCGCTTCTGTCGCGGTGCGGACGGGACTGCGGAG 1020
Db 961 TCCGTTATTCAGCACTGCTCGCCACCGCTTCTGTCGCGGTGCGGACGGGACTGCGGAG 1020
Qy 1021 CTCACCAACCGCTGCTACCGCTTATGAAGGACCTCTATTTACTAGTACTAATGCT 1080
Db 1021 CTCACCAACCGCTGCTACCGCTTATGAAGGACCTCTATTTACTAGTACTAATGCT 1080
Qy 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCTGTTTCAACCTTCTGACACTCTGCTT 1140
Db 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCTGTTTCAACCTTCTGACACTCTGCTT 1140
Qy 1141 GCGCGCTGCGCGACAGAAATTTGATTTCTGCGGTGGTGGCAGCTGTTCTACCTCCGTC 1200
Db 1141 GCGCGCTGCGCGACAGAAATTTGATTTCTGCGGTGGTGGCAGCTGTTCTACCTCCGTC 1200
Qy 1201 GTTGTCTAGCAATGCGCGAGCGGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db 1201 GTTGTCTAGCAATGCGCGAGCGGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
Qy 1261 CAGGATAGGGTATTCGAATCCCGATGACATTCGCGGAGATCTCGTGTGTTAT 1320
Db 1261 CAGGATAGGGTATTCGAATCCCGATGACATTCGCGGAGATCTCGTGTGTTAT 1320
Qy 1321 CAGGATATGATTAACCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1380
Db 1321 CAGGATATGATTAACCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1380
```

[illegible]

Db 1126 GTGCGTGAGATCGGCGCGGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1185
Qy 1141 GCGGCGCTCCGACAGAAATGATTTCTGTCGGCTGGTGGCCAGCTGTCTACTCCGTCCTC 1200
Db 1186 GCGGCGCTCCGACAGAAATGATTTCTGTCGGCTGGTGGCCAGCTGTCTACTCCGTCCTC 1245
Qy 1201 GTTGTCTCAGCCCAATGGCGAGCGAGCTGTTAAAGTTGTATATCATCTGTAGAGAAATGCTCAG 1260
Db 1246 GTTGTCTCAGCCCAATGGCGAGCGAGCTGTTAAAGTTGTATATCATCTGTAGAGAAATGCTCAG 1305
Qy 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGGTTATT 1320
Db 1306 CAGGATAAGGGTATTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGGTTATT 1365
Qy 1321 CAGGATTATGATACCAACATGAACAGATCGCGCAGCGCTTCTCCAGCCCATCGCGC 1380
Db 1366 CAGGATTATGATACCAACATGAACAGATCGCGCAGCGCTTCTCCAGCCCATCGCGC 1425
Qy 1381 CTTTCTCTGCTCTCGAGCTAATGATGTGCTTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
Db 1426 CTTTCTCTGCTCTCGAGCTAATGATGTGCTTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1485
Qy 1441 GACCACTGCATTTAGCTCTTCTGACCTGGCCAGTTATGTTTCTGACTCTGTGACCTTG 1500
Db 1486 GACCACTGCATTTAGCTCTTCTGACCTGGCCAGTTATGTTTCTGACTCTGTGACCTTG 1545
Qy 1501 GTTAAATGTTGGACCGCGCGAGCGGCTGTTGCCGGTTCGCTGATTTGGACCAAGGTAC 1560
Db 1546 GTTAAATGTTGGACCGCGCGAGCGGCTGTTGCCGGTTCGCTGATTTGGACCAAGGTAC 1605
Qy 1561 CTTGACGGTGCCTCTCTCCACATCCAGCAGTACTCGAAGACCTTTCTTTGCTGCGG 1620
Db 1606 CTTGACGGTGCCTCTCTCCACATCCAGCAGTACTCGAAGACCTTTCTTTGCTGCGG 1665
Qy 1621 CTCGCGGTAAAGCTCTCTTTCTGGAGGAGGAGGACCAACTAAAGCCGGGTACCTTATAAT 1680
Db 1666 CTCGCGGTAAAGCTCTCTTTCTGGAGGAGGAGGACCAACTAAAGCCGGGTACCTTATAAT 1725
Qy 1681 TATAACACCACTGCTAGGACCACTGCTGTCGAGAAATGCGCGCGGACCGGTCGCT 1740
Db 1726 TATAACACCACTGCTAGGACCACTGCTGTCGAGAAATGCGCGCGGACCGGTCGCT 1785
Qy 1741 ATTTCACCTTACACCACTAGCCTGGTGTGCTGCTCCCATTTCTGCGGTTCGCGTT 1800
Db 1786 ATTTCACCTTACACCACTAGCCTGGTGTGCTGCTCCCATTTCTGCGGTTCGCGTT 1845
Qy 1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTGAGGATACCTTGACCTACCTGCGCGGCG 1860
Db 1846 TTAGCCCCCACTCTGCGCTAGCAATGCTTGAGGATACCTTGACCTACCTGCGCGGCG 1905
Qy 1861 CATACCTTTGATGATTTCTGCCAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
Db 1906 CATACCTTTGATGATTTCTGCCAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1965
Qy 1921 CAGTCTACTGCTGCTGAGCTTCAAGCGCTTTAAGATGAAGGTGGGTAAACTCGGGAGTTG 1980
Db 1966 CAGTCTACTGCTGCTGAGCTTCAAGCGCTTTAAGATGAAGGTGGGTAAACTCGGGAGTTG 2025
Qy 1981 TAG 1983
Db 2026 TAG 2028

RESULT 6
AR232463
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AR232463
Sequence 1 from patent US 6455492.
AR232463
AR232463.1 GI:27274532
Unknown.
Unknown.
Unclassified.

2094 bp mRNA linear PAT 20-DEC-2002

REFERENCE 1 (bases 1 to 2094)
AUTHORS Reyes, G.R., Bradley, D.W., Twu, J.-S., Purdy, M.A., Tam, A.W.,
Krawczynski, K.Z. and Yarbough, P.O.
TITLE Hepatitis E virus vaccine and method
JOURNAL Patent: US 6455492-A 1 24-SEP-2002;
FEATURES Location/Qualifiers
source 1..2094
/organism="unknown"
/mol_type="mRNA"
ORIGIN
Query Match 100.0%; Score 1983; DB 6; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCGCGCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGTCGCCGGGCCA 60
Db 46 ATGCGCGCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGTCGCCGGGCCA 105
Qy 61 CGCGCGGTCAGCGCTGCGCGCGCTGCTGCGCGCGGCGAGCGGCTTCCGGCGGTGT 120
Db 106 CGCGCGGTCAGCGCTGCGCGCGCTGCTGCGCGCGGCGAGCGGCTTCCGGCGGTGT 155
Qy 121 TTCTGGGCTGACCGGCTTGAATCTCAGCCCTTTCGCAATCCCTATATTCATCCAAACA 180
Db 166 TTCTGGGCTGACCGGCTTGAATCTCAGCCCTTTCGCAATCCCTATATTCATCCAAACA 225
Qy 181 CCCTTCGCGCGCATGTCACCGCTGCGCGCGGCTGAGCTGCTGCTGCGCAACCGGCC 240
Db 226 CCCTTCGCGCGCATGTCACCGCTGCGCGCGGCTGAGCTGCTGCTGCGCAACCGGCC 285
Qy 241 CGACCACTCGGCTCGGCTGCGCGCGGCTGAGCGCGCGCGCGCTGCTGCTGCGCTCGT 300
Db 286 CGACCACTCGGCTCGGCTGCGCGCGGCTGAGCGCGCGCGCGCTGCTGCTGCGCTCGT 345
Qy 301 AGACCTTACCACTCGGCGCGCGCTGAGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCTG 360
Db 346 AGACCTTACCACTCGGCGCGCGCTGAGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCTG 405
Qy 361 CGAGTGCCTGATGTCGACTCCGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTG 420
Db 406 CGAGTGCCTGATGTCGACTCCGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTG 465
Qy 421 TCTCCCTTACCTTTCGCTGCGCGCGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 466 TCTCCCTTACCTTTCGCTGCGCGCGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 525
Qy 481 AGTCCGCTTTTACCTTTCGCTGCGCGCGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 526 AGTCCGCTTTTACCTTTCGCTGCGCGCGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCT 585
Qy 541 AATTATGCCAGTACCGGCTGCGCGCGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 586 AATTATGCCAGTACCGGCTGCGCGCGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
Qy 601 GCTGTCGCGGTTTACCGCATCTCCATCTCATCTGCGCGCGCTGAGCGCGCTGCTGCTGCTG 660
Db 646 GCTGTCGCGGTTTACCGCATCTCCATCTCATCTGCGCGCGCTGAGCGCGCTGCTGCTGCTG 705
Qy 661 TCGGTTGATATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
Db 706 TCGGTTGATATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 765
Qy 721 GCTCTGAGCTTGTGATTCCTCAAGTACGCGCTTACCTATCTGTAACCAAGGCTGCGCTCC 780
Db 766 GCTCTGAGCTTGTGATTCCTCAAGTACGCGCTTACCTATCTGTAACCAAGGCTGCGCTCC 825
Qy 781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 826 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTG 885
Qy 841 CATGGCTACTCGTAAATTCCTATATACTAATAACCTATACCGGCTGCTGCTGCTGCTGCTG 900

886 CATGGCTCACTCGTAAATTCCTACTATAATACACCTATACCGGTGCCCTCGGGCTGTG 945
901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTC 960
946 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTC 1005
961 TCCGGTTATTCCAGCACTGCTCGCCACCGCTTCGTGCGGTGGAGCGGACTGCGGAG 1020
1006 TCCGGTTATTCCAGCACTGCTCGCCACCGCTTCGTGCGGTGGAGCGGACTGCGGAG 1065
1021 CTCACCAACCGGCTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGT 1080
1066 CTCACCAACCGGCTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGT 1125
1081 GTGGTGAGATGCGCGGGGATAGCCCTACCTGTTCACCTTGTCAACCTTGTGACACTCTGCT 1140
1126 GTGGTGAGATGCGCGGGGATAGCCCTACCTGTTCACCTTGTTCACCTTGTGACACTCTGCT 1185
1141 GCGGGCTGCGGACAGAAATTGATTTGTCGGCTGGTGGCCAGCTGTTTCTACTCCCGTCCC 1200
1186 GCGGGCTGCGGACAGAAATTGATTTGTCGGCTGGTGGCCAGCTGTTTCTACTCCCGTCCC 1245
1201 GTTGTCTCAGCAATGCGGACCGCACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
1246 GTTGTCTCAGCAATGCGGACCGCACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1305
1261 CAGGATTAAGGTATTGCAATCCCGCAGTACATTCACCTCGGAGAACTCTGTTGTTATT 1320
1306 CAGGATTAAGGTATTGCAATCCCGCAGTACATTCACCTCGGAGAACTCTGTTGTTATT 1365
1321 CAGGATTAATGAACCAACATGAACAGATCGGCGGACGCTTCTCAGGCCCACTCGGC 1380
1366 CAGGATTAATGAACCAACATGAACAGATCGGCGGACGCTTCTCAGGCCCACTCGGC 1425
1381 CTTTCTCTGTCCTTCGAGCTAATGATGTGCTTTGGTCTCTCTCAACGCTGCGGAGTAT 1440
1426 CTTTCTCTGTCCTTCGAGCTAATGATGTGCTTTGGTCTCTCTCAACGCTGCGGAGTAT 1485
1441 GACAGTCCACTTATGCTCTTCACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
1486 GACAGTCCACTTATGCTCTTCACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1545
1501 GTTAATGTTGGACCGGCGGACGCGGTTCGCCGCTCGCTCGATTGGACCAAGGTACA 1560
1546 GTTAATGTTGGACCGGCGGACGCGGTTCGCCGCTCGCTCGATTGGACCAAGGTACA 1605
1561 CTTGACGGTCCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCGG 1620
1606 CTTGACGGTCCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCGG 1665
1621 CTCGCGGTAGCTCTCTTTCTGGGAGCGGACCACTAAGCGGGTACCTTATAAT 1680
1666 CTCGCGGTAGCTCTCTTTCTGGGAGCGGACCACTAAGCGGGTACCTTATAAT 1725
1681 TATAACCACTAGCTAGCGACCACTGCTTGTGAGAAATCGCGCGGACACCGGGTGCCT 1740
1726 TATAACCACTAGCTAGCGACCACTGCTTGTGAGAAATCGCGCGGACACCGGGTGCCT 1785
1741 ATTTCCACTTACACCACTAGCTAGCGACCACTGCTTGTGAGAAATCGCGGGTGGCGTT 1800
1786 ATTTCCACTTACACCACTAGCTAGCGACCACTGCTTGTGAGAAATCGCGGGTGGCGTT 1845
1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTCAGAGATACCTTGGACTACCTGCCCGGCC 1860
1846 TTAGCCCCCACTCTGCGCTAGCAATGCTTCAGAGATACCTTGGACTACCTGCCCGGCC 1905
1861 CATACTTTTGTATGATTTCTGCGCTAGCAATGCTTCAGAGATACCTTGGACTACCTGCCCGGCC 1920
1906 CATACTTTTGTATGATTTCTGCGCTAGCAATGCTTCAGAGATACCTTGGACTACCTGCCCGGCC 1965
1921 CAGTCTACTGTCGCTGAGCTTCAAGGTGAGTGAAGGTGGGTAAACCTCGGGAGTTG 1980
1966 CAGTCTACTGTCGCTGAGCTTCAAGGTGAGTGAAGGTGGGTAAACCTCGGGAGTTG 2025

Qy 1981 TAG 1983
Db 2026 TAG 2028

RESULT 7
LOCUS ARL150915 7195 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 6 from patent US 6229005.
ACCESSION ARL150915
VERSION ARL150915.1 GI:15115506
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7195)
AUTHORS Reyes,G.R., Yarbough,P.O., Bradley,D.W., Krawczynski,K.Z., Tam,A.W.
and Fry,K.E.
TITLE DNA sequences of enterically transmitted non-A/non-B hepatitis
viral agent
JOURNAL Patent: US 6229005-A 6 08-MAY-2001;
FEATURES Location/Qualifiers
source 1..7195
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 1983; DB 6; Length 7195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTTCCTATGCTGCGCGCGCA 60
Db 5147 ATGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTTCCTATGCTGCGCGCGCA 5206
Qy 61 CCGCGCCGTCAGCGCTCGCGCGCGCTGCGGGCGGCGAGCGGCGGTTCGCGGCGTGT 120
Db 5207 CCGCGCCGTCAGCGCTCGCGCGCGCTGCGGGCGGCGAGCGGCGGTTCGCGGCGTGT 5266
Qy 121 TTTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATTCCTATATTCATCCAAAC 180
Db 5267 TTTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATTCCTATATTCATCCAAAC 5326
Qy 181 CCTTTCGCCCGGATGTCACGCTGCGCGCGGCTGAGCTCTGTTTCGCAACCGGCC 240
Db 5327 CCTTTCGCCCGGATGTCACGCTGCGCGCGGCTGAGCTCTGTTTCGCAACCGGCC 5386
Qy 241 CGACCACTCGGCTCGGCTTGGCGTGACCGCGCGCGCGCGCGCGCTTGCCTCAGTGT 300
Db 5387 CGACCACTCGGCTCGGCTTGGCGTGACCGCGCGCGCGCGCGCTTGCCTCAGTGT 5446
Qy 301 AGACCTACCACTGGGGCGCGCGCTAACCGCGGTGCTCCGGGCCATGACACCCG 360
Db 5447 AGACCTACCACTGGGGCGCGCGCTAACCGCGGTGCTCCGGGCCATGACACCCG 5506
Qy 361 CCAGTGCTCGATGTCGACTCCCGCGCGCATCTTGGCGCGGAGTAACTTATCAACA 420
Db 5507 CCAGTGCTCGATGTCGACTCCCGCGCGCATCTTGGCGCGGAGTAACTTATCAACA 5566
Qy 421 TCTCCCTTACCTCTCCGTGGCCACCGGCACTAACCTGGTTCCTTATGCGGCCCTCT 480
Db 5567 TCTCCCTTACCTCTCCGTGGCCACCGGCACTAACCTGGTTCCTTATGCGGCCCTCT 5626
Qy 481 AGTCGGTTTTTACCCCTTCAGGACGGCACCAATACCATATATGCGCCAGGACTTCT 540
Db 5627 AGTCGGTTTTTACCCCTTCAGGACGGCACCAATACCATATATGCGCCAGGACTTCT 5686
Qy 541 AATTATGCCAGTACCGGGTTGCCGTGCGCAATTCGTTTACCGCCCGCTGGTCCCCAAT 600
Db 5687 AATTATGCCAGTACCGGGTTGCCGTGCGCAATTCGTTTACCGCCCGCTGGTCCCCAAT 5746
Qy 601 GCTGTCGGCGGTTACGGCAATCTCCATCTCATTTCTTGGGCCACAGACCAACCCCGGAG 660

|||||
5747 GCTGTGGCGGTAGCCCATCTCCATCTCATTTCTGGCCACAGACCACACCCCGAGC 5806
QY |||||
661 TCCGTTGATATGAATTAATTAACCTTCGACGGATGTTGTAATTTAGTCAGCCCGGATA 720
Db |||||
5807 TCCGTTGATATGAATTAATTAACCTTCGACGGATGTTGTAATTTAGTCAGCCCGGATA 5866
QY |||||
721 GCCTCTGAGCTTGATATCCCAAGTCAGCGCTACACTATCGTAACCAAGGCTGGCGCTCC 780
Db |||||
5867 GCCTCTGAGCTTGATATCCCAAGTCAGCGCTACACTATCGTAACCAAGGCTGGCGCTCC 5926
QY |||||
781 GTCGAGACTCTTGGGGTGGCTGAGGAGGAGGCTACCTCTGTGCTTGTATGCTTTGCATA 840
Db |||||
5927 GTCGAGACTCTTGGGGTGGCTGAGGAGGAGGCTACCTCTGTGCTTGTATGCTTTGCATA 5986
QY |||||
841 CATGCTCACTGTAATAATTCCTATATAACACCTATACCGGTATACCGGTGCGCTGGGCTGTTG 900
Db |||||
5987 CATGGCTCACTGTAATAATTCCTATATAACACCTATACCGGTATACCGGTGCGCTGGGCTGTTG 6046
QY |||||
901 GACTTTGCCCTTGAGCTTGAGTTTTCGCAACCTTACCCCGGTAAACACCAATACGCGGGTC 960
Db |||||
6047 GACTTTGCCCTTGAGCTTGAGTTTTCGCAACCTTACCCCGGTAAACACCAATACGCGGGTC 6106
QY |||||
961 TCCCGTTATTCAGCACTGCTGCCACCGCCTTCGTCGGGTGCGGACGGGACTGCCGAG 1020
Db |||||
6107 TCCCGTTATTCAGCACTGCTGCCACCGCCTTCGTCGGGTGCGGACGGGACTGCCGAG 6166
QY |||||
1021 CTCACACACAGCGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGTT 1080
Db |||||
6167 CTCACACACAGCGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGTT 6226
QY |||||
1081 GTCGCTGAGATCGCGCGGGATAGCCCTCACCTCTGTTCAACCTTGCTGACACTCTGTT 1140
Db |||||
6227 GTCGCTGAGATCGCGCGGGATAGCCCTCACCTCTGTTCAACCTTGCTGACACTCTGTT 6286
QY |||||
1141 GCGCGCCTGCCACAGAAATTTGATTTGTCGGCTGTTGGCAGCTGTTTACTCCCGTCCC 1200
Db |||||
6287 GCGCGCCTGCCACAGAAATTTGATTTGTCGGCTGTTGGCAGCTGTTTACTCCCGTCCC 6346
QY |||||
1201 GTTGTCTAGCAATGGCGAGCGGACTGTTAGTTGTAATACATCTGTAGAGAAATGCTCAG 1260
Db |||||
6347 GTTGTCTAGCAATGGCGAGCGGACTGTTAGTTGTAATACATCTGTAGAGAAATGCTCAG 6406
QY |||||
1261 CAGGATAGGGTATTGCAATCCCGATGACATGACCTCGAGAAATCTCGTGGTTATT 1320
Db |||||
6407 CAGGATAGGGTATTGCAATCCCGATGACATGACCTCGAGAAATCTCGTGGTTATT 6466
QY |||||
1321 CAGGATTTATGATAACCAACATGAACAGATCGGCGGACGCTTCTCCAGCCCAATCGCG 1380
Db |||||
6467 CAGGATTTATGATAACCAACATGAACAGATCGGCGGACGCTTCTCCAGCCCAATCGCG 6526
QY |||||
1381 CTTTCTCTGCTGCTGAGCTAATGATGCTTTGGCTCTCTCTACCGCTGCGAGTAT 1440
Db |||||
6527 CTTTCTCTGCTGCTGAGCTAATGATGCTTTGGCTCTCTCTACCGCTGCGAGTAT 6586
QY |||||
1441 GACAGTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGACCTTG 1500
Db |||||
6587 GACAGTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGACCTTG 6646
QY |||||
1501 GTTAAATGTTGCGACCGCGCGAGGCGGTTGCCCGGTGCTCGATTGGACCAAGGTCA 1560
Db |||||
6647 GTTAAATGTTGCGACCGCGCGAGGCGGTTGCCCGGTGCTCGATTGGACCAAGGTCA 6706
QY |||||
1561 CTTGACGGTGGCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTTGTCTGGCG 1620
Db |||||
6707 CTTGACGGTGGCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTTGTCTGGCG 6766
QY |||||
1621 CTTCCGGTTAAGCTCTCTTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCCCTTATAAT 1680
Db |||||
6767 CTTCCGGTTAAGCTCTCTTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCCCTTATAAT 6826
QY |||||
1681 TATAACACCACTGCTAGCAGCAACTGCTTGTTCGAGAATGCCCGCGGACCGGGTCTGCT 1740
|||||

Db 6827 TATAACCACTGCTAGCAGCAACTGCTTGTTCGAGAATGCCCGGACCGGGTCTGCT 6886
QY 1741 ATTTCACCTTACACCACTAGCTGGGTGCTGTGTCCTGCTCTCCATTTCTGCGGTTCGCGTT 1800
Db 6887 ATTTCACCTTACACCACTAGCTGGGTGCTGTGTCCTGCTCTCCATTTCTGCGGTTCGCGTT 6946
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTGCTTGAGGATACCTTGGACTACCTCCCGCGGCC 1860
Db 6947 TTAGCCCCCACTCTGCGCTAGCATTGCTTGAGGATACCTTGGACTACCTCCCGCGGCC 7006
QY 1861 CATACTTTTGTATGATTTCTGCCAGAGTCCGCCCTTGGCTTTCAGGGCTGCGCTTTC 1920
Db 7007 CATACTTTTGTATGATTTCTGCCAGAGTCCGCCCTTGGCTTTCAGGGCTGCGCTTTC 7066
QY 1921 CAGTCTACTGCTGAGCTTACAGCTTACAGCTTAAAGTGAAGTGGTAAACTCGGAGTTG 1980
Db 7067 CAGTCTACTGCTGAGCTTACAGCTTAAAGTGAAGTGGTAAACTCGGAGTTG 7126
QY 1981 TAG 1983
Db 7127 TAG 7129

RESULT 8
AR207633
LOCUS AR207633 7195 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 6 from patent US 6379891.
ACCESSION AR207633
VERSION AR207633.1 GI:21507438
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7195)
AUTHORS Reyes,G.R.; Yarbrough,P.O.; Bradley,D.W.; Krawczynski,K.Z.; Tam,A.
and Fry,K.E.
TITLE Method of detecting HEV infection
JOURNAL Patent: US 6379891-A 6 30-APR-2002;
FEATURES
source 1..7195
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 1983; DB 6; Length 7195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCGGCCCTATTTTGTGCTGCTCTCATGTTTTCCTATGCTGCCCGGCCA 60
Db 5147 ATGCGCCCTCGGCCCTATTTTGTGCTGCTCTCATGTTTTCCTATGCTGCCCGGCCA 5206
QY 61 CGCCCGGTACGCGTCTGGCCCGCTGTGGGCGCGCAGCGGCTTCGGCGGTGTT 120
Db 5207 CGCCCGGTACGCGTCTGGCCCGCTGTGGGCGCGCAGCGGCTTCGGCGGTGTT 5266
QY 121 TTCTGGGTGACGGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACAAC 180
Db 5267 TTCTGGGTGACGGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACAAC 5326
QY 181 CCCTTCGCCCGCCGATGTCAACCGCTGCGCGCGGGCTGGACCTCGTGTTCGCAACCCGCC 240
Db 5327 CCCTTCGCCCGCCGATGTCAACCGCTGCGCGCGGGCTGGACCTCGTGTTCGCAACCCGCC 5386
QY 241 CGACCACTCGGTCTCGCTTGGGTGACAGGCGCCAGCGCCCGCTTGCCTTCACTGTT 300
Db 5387 CGACCACTCGGTCTCGCTTGGGTGACAGGCGCCAGCGCCCGCTTGCCTTCACTGTT 5446
QY 301 AGACCTTACCACTGGGGCGCGCTAAACCGGGTTCGCTCCGCCCATGACACCCCG 360
Db 5447 AGACCTTACCACTGGGGCGCGCTAAACCGGGTTCGCTCCGCCCATGACACCCCG 5506
QY 361 CCAGTGCCTGATGTCGACTCCCGCGCGCCATCTTTGCGCGGAGTATAACCTATCAACA 420

Db 5507 CCAAGTGCCTGATGTGCAGTCCCGCGGCGCATCTTGGCCGGCAGTATAACCTATCAACA 5566
Qy 421 TCTCCCTTACCTCTTCGCGTGGGCAACCGGCACTAAACCTGTGTTCTTTATGCGGCCCTCTT 480
Db 5567 TCTCCCTTACCTCTTCGCGTGGGCAACCGGCACTAAACCTGTGTTCTTTATGCGGCCCTCTT 5626
Qy 481 AGTCGGCTTTACCCCTTCAGGAGCGGCAACCAATACCATATAATATGGCCACGGAGCTTCT 540
Db 5627 AGTCGGCTTTACCCCTTCAGGAGCGGCAACCAATACCATATAATATGGCCACGGAGCTTCT 5686
Qy 541 AATTATGCCAGTACCGGGTTCGCCGTGCCACCAATCCGTTACCGCCCGCTGGTCCCAAT 600
Db 5687 AATTATGCCAGTACCGGGTTCGCCGTGCCACCAATCCGTTACCGCCCGCTGGTCCCAAT 5746
Qy 601 GCTGTGGCGGGTACGGCATCTCCATCTCACTATCTTCTGGCCACAGACCAACCCCGAGC 660
Db 5747 GCTGTGGCGGGTACGGCATCTCCATCTCACTATCTTCTGGCCACAGACCAACCCCGAGC 5806
Qy 661 TCCGTTGATATGAATCAATAACCTCGACGGATGTTTCGTAATTTTAGTCCAGCCCGGCATA 720
Db 5807 TCCGTTGATATGAATCAATAACCTCGACGGATGTTTCGTAATTTTAGTCCAGCCCGGCATA 5866
Qy 721 GCCTCTGAGCTTGATGCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGTCC 780
Db 5867 GCCTCTGAGCTTGATGCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGTCC 5926
Qy 781 GTCCAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCGTCTGTGTTATGCTTTGCATA 840
Db 5927 GTCCAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCGTCTGTGTTATGCTTTGCATA 5986
Qy 841 CATGGCTCACTCGTAATTCCTTACTACTAATAACACCTTATACCGGTGCCCTCGGGCTGTTG 900
Db 5987 CATGGCTCACTCGTAATTCCTTACTACTAATAACACCTTATACCGGTGCCCTCGGGCTGTTG 6046
Qy 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATAACCGGGTC 960
Db 6047 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATAACCGGGTC 6106
Qy 961 TCCGGTATTCAGCACTGCTCGCCACCGCTTGTGCGGGTGGGACGGAGCTGGCAG 1020
Db 6107 TCCGGTATTCAGCACTGCTCGCCACCGCTTGTGCGGGTGGGACGGAGCTGGCAG 6166
Qy 1021 CTCACCAACCGCTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGT 1080
Db 6167 CTCACCAACCGCTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGT 6226
Qy 1081 GTCGGTGAGATCGGCCCGGGATAGCCCTCACCTGTTCACCTTGCTGACACCTCTGCTT 1140
Db 6227 GTCGGTGAGATCGGCCCGGGATAGCCCTCACCTGTTCACCTTGCTGACACCTCTGCTT 6286
Qy 1141 GCGGGCTGCGGACAGAAATGATTCGTGGCTGGTGGCCAGCTGTTCTACTCCGCTCC 1200
Db 6287 GCGGGCTGCGGACAGAAATGATTCGTGGCTGGTGGCCAGCTGTTCTACTCCGCTCC 6346
Qy 1201 GTTGTCTCAGCAATGCGGACCGCAGTGTAAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db 6347 GTTGTCTCAGCAATGCGGACCGCAGTGTAAAGTTGTATACATCTGTAGAGAATGCTCAG 6406
Qy 1261 CAGGATAAGGGTATTCGAATCCCGCATGACATTTGAGTGTATACATCTGTAGAGAATGCTCAG 1320
Db 6407 CAGGATAAGGGTATTCGAATCCCGCATGACATTTGAGTGTATACATCTGTAGAGAATGCTCAG 6466
Qy 1321 CAGGATTATGATAACCAACATGAACAGATCGGCCAGCGCTTCTCCAGCCCATCGGC 1380
Db 6467 CAGGATTATGATAACCAACATGAACAGATCGGCCAGCGCTTCTCCAGCCCATCGGC 6526
Qy 1381 CCTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCAACCGCTGCCGAGTAT 1440
Db 6527 CCTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCAACCGCTGCCGAGTAT 6586
Qy 1441 GACAGTCCACTTATGGCTCTTGA CTGGCCAGTTATGTTTCTGACTGTGACCTTG 1500

Db 6587 GACCAGTCCACTTATGGCTCTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 6646
Qy 1501 GTTAAATGTTGAGACCGCGCGAGGGCGTTGCGCGGTGCTCGATTGGACAAGGTCAACA 1560
Db 6647 GTTAAATGTTGAGACCGCGCGAGGGCGTTGCGCGGTGCTCGATTGGACAAGGTCAACA 6706
Qy 1561 CTTGAGCGTCCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTGTCCTCGCG 1620
Db 6707 CTTGAGCGTCCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTGTCCTCGCG 6766
Qy 1621 CTCGCGGTAAGCTCTCTTTCTCGGAGGCGAGGCAACAATAAAGCCGGGTACCTTATAAT 1680
Db 6767 CTCGCGGTAAGCTCTCTTTCTCGGAGGCGAGGCAACAATAAAGCCGGGTACCTTATAAT 6826
Qy 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATCGCGCGGCAACCGGGTGCCT 1740
Db 6827 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATCGCGCGGCAACCGGGTGCCT 6886
Qy 1741 ATTTCACATTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTCGCGTTCGCGTT 1800
Db 6887 ATTTCACATTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTCGCGTTCGCGTT 6946
Qy 1801 TTAGCCCCCCTCTCGCTAGCATTCTTGAGGATACCTTGGAATAACCTGCCCGCGCC 1860
Db 6947 TTAGCCCCCCTCTCGCTAGCATTCTTGAGGATACCTTGGAATAACCTGCCCGCGCC 7006
Qy 1861 CATACTTTGATGATTTGCCCCAGAGTGGCGGCCCTTGGCCCTCAGGCTCGGCTTTC 1920
Db 7007 CATACTTTGATGATTTGCCCCAGAGTGGCGGCCCTTGGCCCTCAGGCTCGGCTTTC 7066
Qy 1921 CAGTCTACTGTCCTCAGCTTTCAGCGCTTAAAGTGAAGTGGGTAAACTCGGGAGTTG 1980
Db 7067 CAGTCTACTGTCCTCAGCTTTCAGCGCTTAAAGTGAAGTGGGTAAACTCGGGAGTTG 7126
Qy 1981 TAG 1983
Db 7127 TAG 7129

RESULT 9
HPESVP
LOCUS
DEFINITION
HPESVP
Hepatitis E structural and nonstructural protein mRNA, complete
cds'.
ACCESSION
M73218.1
VERSION
M73218.1
KEYWORDS
nonstructural protein; viral structural protein.
SOURCE
Hepatitis E virus
ORGANISM
Hepatitis E virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Hepatitis E-like viruses.
REFERENCE
1 (bases 1 to 7207)
AUTHORS
Tam,A.W., Smith,M.M., Guerra,M.E., Huang,C.C., Bradley,D.W., Fry,K.E. and Reyes,G.R.
TITLE
Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length viral genome
JOURNAL
Virology 185 (1), 120-131 (1991)
MEDLINE
92024067
PubMed
1926770
COMMENT
Original
source text: Hepatitis E virus cDNA to genomic RNA.
Location/Qualifiers
1..7207
/organism="Hepatitis E virus"
/mol_type="genomic RNA"
/db_xref="taxon:12461"
28..5109
/note="putative"
/codon_start=1
/product="nonstructural protein"
/protein_id="AA045734.1"
/db_xref="GI:330024"
1..7207
/organism="Hepatitis E virus"
/mol_type="genomic RNA"
/db_xref="taxon:12461"
28..5109
/note="putative"
/codon_start=1
/product="nonstructural protein"
/protein_id="AA045734.1"
/db_xref="GI:330024"

VHRCFLRPVGRDVRWYATPTTRGPAANCRRSALRGLPAADRTYCLDGFSGCNFPBAETG
IALYSLHOMSPDVAEAMFRHCHWTRLVYALHLPPEVLLPPGTYRTASVLLIHDGRVV
VYIEGDTNAGYHNDVNLRSWTRTKVTCGDHDLVIERVRAIGHCHVILLITAAPEBSPM
PVPYPRSTEVYTRISIFGGGPTSLFPTSCSTCSRTFHVPAHAIWDRMLFGAIDLOQA
PCSLIMTYLRIGSYKVTYVITWNEGMNASEDALTAAYLITACHORYLRTQALIS
KMRRLERHAQKFTIRLYSWIENGECECTFLOLEFYAOCRRWLSAGFHLDPRLV
PESALCHDRTAIRALKSFCCFMKWLGOECTFLOLEFYAOCRRWLSAGFHLDPRLV
PESALSDISGVYVPGTALQPLQALDPAEIVARAGRLTATVKVSOVDYEDCETL
LGNKTRFSFVGALENGPERNLSPDASQTAAGPFLSYTAASAGLEVRVVA
GLDHRVAPGSPSAGEVTAFCALYRFNREARQHSLLIGNLHFEPLGLLFAFP
SGHWNESANPFCSESTLYRTWSEVDVSSPARDLGFMSEPSIPSRRAATPLAALP
PPAPDPPSPAPALAEPAAGACAPAIHQTAHRRLLETFYDPSKVFAGSLES
TCTWLVNANVDRPGGLCHAFYQYRYPASFDAAQFVWRDGAAYTLTPRIHIAVAP
DYLRHNKRLAEARETCSRLGTAAYPLGLTYGVPTGPSFADAWERNHRPGDELXL
PDLARPEANPTEPTLTITEDVARTANLAJELDSATDVGRACGRVTPGVYOYF
TAGVFGSKRSITQADVVDVVPTRELNWRNRGFAAFPHTAARVYQGRVVIDE
APSLPHLLHLLHAGTAVHLLGDPNQIPADIFEHAGIVPAIRPDLPSTVMMHWTRW
PADVCELRGLGPMYQTSRLVRSFWGEPVQKLVFTQAAKPNVGSVTHAEQA
TYTETIITADARGLIQSSRAHAIVALTRHTEKVIIDAPGLLREVGISDAIVNFF
LAGGETIHORPSVIRPGRNDPANDVTLAAPPSCQISAPHOLAEELGHRPVPVAALVP
CPLEBGLLYLQELTTCDSVUTEFELTDIVHCRMAAPQSKAVLTLVGRYGRTKLY
NASHSDRSLARFIPAIGPVQVTTCELYELVEAMVEKQDGSVLELDCNRDSRI
TFQDKCNKFTGETIAHKVGVGSAMSKTFCALFGPMFRAIEKAILLALDQGVYFG
DAFDOTFSAVAAKASAMFENDEFSDTQNFSLGLECAIMECECGPQWLIRLYH
LIRSAWILQAPKESIRGFKWHSBPGTLLNVTNMVAVITHCYDFRDPQVAAFGRDD
STVLCEYRQSPGAALVJAGCGLKLKVDPRPGLVAGVVVAPGLGALPDVRFAGRLT
EKNWGPSPRAQLBLAVSDFLRKLTVNAQMCVDVSVRYGVSPGLVHNLIGMLQAVA
DGAHPTESVSKPVLDTNLSILCRVE"
5106..5477
/codon_start=1
/product="structural viral protein"
/protein_id="AA45735.1"
/db_xref="GI:330025"
translation="MNMSFAAPMGSRPCALGLFCCSCFCLCPHRPVRSLAAV
GAAAVPAVSGVTGLIUSPSQSPFIQTPSPMSPURPGLDLVFNANPDHSPALVG
TRSPAPPLPHVVDLPQLGRPR"
5147..7129
/codon_start=1
/product="structural viral protein"
/protein_id="AA45736.1"
/db_xref="GI:330026"
translation="NRPRILLLLMLPMLPAPPQSPQSRRRRRSGSGGFGWGD
RVDSOPFAIPIHTPNPPADVPDVAAGRPVRQAPPLGSAWRDQAPAVASRRRP
TTAGRAPLTAAPADPTPVDVDSRGAILRQYNLSTSPLTSSVATQNLVLYAALP
SELLPDGTNTHIMATASNAQIRVARATIRPLVNAVGGYALISFNPQTITI
PISVDNLSITSDRILVQPIASELVIPEERLHYNQWRSVETSGVAREATSLGV
MLCIGSLVSNYNTPTGALGLDPALEFNLTPGNTNTRVSRYSSTARHSLRG
ADGTAEALTTAATRFMKDLYFTNGVGEIGRGIALTLFNLADTLGLLPTELSAG
GQLFYSRPVVSANGPVTKLYTSVENAQDQKGIATPHIDLGESRVVIQDYDNQHEQD
RPTSPAPSRPSVLANDVLSLTAAYDQSTYSGSSTGSPVVSVDVTLVNVTAQ
AVARSLDWTKVLDRPPLSTIQYSKTFPVLPLRGKLSFWEAGTTKAGYPVNYNTAS
DQLLVENAGHRAVISTYTSIGAGPVSISAVAVLAPHASALALLEDTLDYPARHTFD
DFPCPRPLGLGQCAQFSTVABELQRLMKVGTREL"

ORIGIN

Query Match 100.0%; Score 1983; DB 14; Length 7207;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCCTATTTTGTGCTGCTCCTCATGTTTTTGGCTATGCTGCCCGGCCA 60
DB 5147 ATGCGCCCTCGGCCTATTTTGTGCTGCTCCTCATGTTTTTGGCTATGCTGCCCGGCCA 5206
QY 61 CGCCCGGCTCAGCGCTCGCGCCCGCTGCTGGCGCGGAGCGCGGTTCGCGCGGTG 120
DB 5207 CGCCCGGCTCAGCGCTCGCGCCCGCTGCTGGCGCGGAGCGCGGTTCGCGCGGTG 5266
QY 121 TTCTGGGGTGACCGGGTGAATCTCAGCCCTTCGCAATCCCTATATTATCCAAAC 180
DB 5267 TTCTGGGGTGACCGGGTGAATCTCAGCCCTTCGCAATCCCTATATTATCCAAAC 5326
QY 181 CCCTTCGCCCCGATGTCACCGCTGCGCGCGGGCTGGACCTCGTGTTCGCCAACCCGCG 240

781 GTCCAGAGCCTCTGGGTGGCTGAGGAGGAGGTACCTCTGGTCTTGTTATGCTTTGCATA 840
 841 CATGGCTCAGTCGTAATTCCTATCTAATTAACACCTATACCGGTGCCCTCGGGCTGTTG 900
 841 CATGGCTCAGCCGTAATTCCTATCTAATTAACACCTATACCGGTGCCCTCGGGCTGTTG 900
 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC 960
 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC 960
 961 TCCCGTTATTCAGCACTGCTCGCCACCGCCCTTCGTGCGGTGCGGACGGGACTGCCGAG 1020
 961 TCCCGTTATTCAGCACTGCTCGCCACCGCCCTTCGTGCGGTGCGGACGGGACTGCCGAG 1020
 1021 CTCACACAGCGCTGCTACCGCTTTATGAAGGACCTTATTTACTAGTACTAATGGT 1080
 1021 CTCACACAGCGCTGCTACCGCTTTATGAAGGACCTTATTTACTAGTACTAATGGT 1080
 1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
 1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
 1141 GSCGCGCTGCCACAGAAATGATTTCTGCGCTGCTGCGGACCTGTTACTTCCCGTCCC 1200
 1141 GSCGCGCTGCCACAGAAATGATTTCTGCGCTGCTGCGGACCTGTTACTTCCCGTCCC 1200
 1201 GTTGTCTAGCAATGCGGAGCGGACTGTTAGTTGTATACATCTGTAGAGAACTCTCAG 1260
 1201 GTTGTCTAGCAATGCGGAGCGGACTGTTAGTTGTATACATCTGTAGAGAACTCTCAG 1260
 1261 CAGGATAGGGTATTGCAATCCGATGATGACCTCGGAGAACTCTGCTGCTGTTAT 1320
 1261 CAGGATAGGGTATTGCAATCCGATGATGACCTCGGAGAACTCTGCTGCTGTTAT 1320
 1321 CAGGATATTATGATACCAACATGAACATCGGCGGACGCTTCTCCAGCCCACTCGCGC 1380
 1321 CAGGATATTATGATACCAACATGAACATCGGCGGACGCTTCTCCAGCCCACTCGCGC 1380
 1381 CTTTCTCTGCTCTGAGCTAATGATGCTTTGGCTCTCTCTCACCCTGCGGATAT 1440
 1381 CTTTCTCTGCTCTGAGCTAATGATGCTTTGGCTCTCTCTCACCCTGCGGATAT 1440
 1441 GACCACTCAGCTTATGCTCTGAGCTGCGGCGGAGCTGTTGTTCTGACCTCTGACCTTG 1500
 1441 GACCACTCAGCTTATGCTCTGAGCTGCGGCGGAGCTGTTGTTCTGACCTCTGACCTTG 1500
 1501 GTTAATGTTGACCGCGCGGAGCGGCTGTCGCGTCTGATTTGACCTGACCTGACCTG 1560
 1501 GTTAATGTTGACCGCGCGGAGCGGCTGTCGCGTCTGATTTGACCTGACCTGACCTG 1560
 1561 CTTGAGCGTGGCGGCTCTGACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGG 1620
 1561 CTTGAGCGTGGCGGCTCTGACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGG 1620
 1621 CTTCCGGGTAGCTCTCTTTCTGGAGGAGGAGCAAACTAAAGCGGGTACCTTATATAT 1680
 1621 CTTCCGGGTAGCTCTCTTTCTGGAGGAGGAGGAGCAAACTAAAGCGGGTACCTTATAT 1680
 1681 TATAACACCACTGCTAGGAGCAAACTGCTGTTGAGAAATGCCCGCGGACCGGGTCTG 1740
 1681 TATAACACCACTGCTAGGAGCAAACTGCTGTTGAGAAATGCCCGCGGACCGGGTCTG 1740
 1741 ATTTCACCTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 1741 ATTTCACCTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 1801 TTAGCCCCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
 1801 TTAGCCCCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
 1861 CATACCTTTGATGATTTCTGCGGAGGAGTGGCGGCGGCTTGGCGCTCAGGGCTGCGCTTC 1920
 1861 CATACCTTTGATGATTTCTGCGGAGGAGTGGCGGCGGCTTGGCGCTCAGGGCTGCGCTTC 1920

QY 1921 CAGTCTACTGTCGTGAGCTTTCAGCGCCTTAAGATGAAGTGGTAAACCTCGGAGTTG 1980
 Db 1921 CAGTCTACTGTCGTGAGCTTTCAGCGCCTTAAGATGAAGTGGTAAACCTCGGAGTTA 1980
 QY 1981 TAG 1983
 Db 1981 TAG 1983
 RESULT 11
 E17109
 LOCUS Hepatitis E virus gene. 7194 bp DNA linear PAT 28-JUL-1999
 DEFINITION Hepatitis E virus gene.
 ACCESSION E17109
 VERSION E17109.1 GI:5711792
 KEYWORDS JP 1998234383-A/5.
 SOURCE unclassified
 ORGANISM unclassified
 1 (baes 1 to 7194)
 Takeda,N., Ri,T. and Miyamura,T.
 HEPATITIS E VIRUS HOLLOW PARTICLE, GENE ENCODING THE SAME AND
 PREPARATION OF RECOMBINATION VECTOR CONTAINING THE GENE AND
 UTILIZATION OF THE RECOMBINATION VECTOR
 Patent: JP 1998234383-A 5 08-SEP-1998;
 DENKA SEIKEN CO LTD, KOKURITSU KANSENSHIYOU KENKYUSHO
 OS Hepatitis E virus
 PN JP 1998234383-A/5
 PD 08-SEP-1998
 PF 28-FEB-1997 JP 1997062445
 PI TAKEDA NAKAZU, RI TENSEI, MIYAMURA TATSUO
 PC C12N15/09,C07K14/08,C12N5/10,C12P21/02;
 CC strandedness: Double;
 CC topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..7194
 FT /organism='Hepatitis E virus'.
 FEATURES
 source
 1..7194
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN
 Query Match 98.3%; Score 1949.4; DB 6; Length 7194;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1962; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
 QY 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCGCGGCA 60
 Db 5147 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCGCGGCA 5206
 QY 61 CGCGCCGCTCAGCCGCTCTGCGCCGCTCTGCGCGCGCGCGCGGCTTCCGCGGTGT 120
 Db 5207 CGCGCCGCTCAGCCGCTCTGCGCCGCTCTGCGCGCGCGCGCGGCTTCCGCGGTGT 5266
 QY 121 TTCTGCGGTGACCGGGTTGATTTCTAGCCCTTCGCGATCCCTATATTTTCATCAACCAAC 180
 Db 5267 TTCTGCGGTGACCGGGTTGATTTCTAGCCCTTCGCGATCCCTATATTTTCATCAACCAAC 5326
 QY 181 CCCTTCGCGCCCGATGTACCGCTCGCGCGCGGGGTGACCTCTGTTTCGCAACCGCGC 240
 Db 5327 CCCTTCGCGCCCGATGTACCGCTCGCGCGCGGGGTGACCTCTGTTTCGCAACCGCGC 5386
 QY 241 CGACCACTCGGCTCGGCTTGGGTGACAGGCGCCAGCGCCCGCGCTTGCCTCAGCTGT 300
 Db 5387 CGACCACTCGGCTCGGCTTGGGTGACAGGCGCCAGCGCCCGCGCTTGCCTCAGCTGT 5446
 QY 301 AGACCTTACCAAGTGGGCGCGCGCTTAAACCGGCTGCTCCGCGCCCATGACACCGCG 360
 Db 5447 AGACCTTACCAAGTGGGCGCGCGCTTAAACCGGCTGCTCCGCGCCCATGACACCGCG 5506

QY	361	CCAGTGCCTGATGTCGACTCCCGCGCGCCATCTTTGGCGGAGTATAACCTATCAACA	420	QY	1441	GACAGTCCACTTATGSGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG	1500
Db	5507	CCAGTGCCTGATGTCGACTCTCGCGGCGCCATCTTTGGCGGAGTATAACCTATCAACA	5566	Db	6587	GACAGTCCACTTATGSGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG	6646
QY	421	TCTCCCTTACCTCTTCCGCGGACCGGACCTAACCTGGTCTCTTTATGCGGCCCTCTTT	480	QY	1501	GTTAAATGTTCCGACCGCGCGCGAGGGCGTTGCGCGGTGCGTCCGATTGGACCAAGGTCA	1560
Db	5567	TCTCCCTTACCTCTTCCGCGGACCGGACCTAACCTGGTCTCTTTATGCGGCCCTCTTT	5626	Db	6647	GTTAAATGTTCCGACCGCGCGCGAGGGCGTTGCGCGGTGCGTCCGATTGGACCAAGGTCA	6706
QY	481	AGTCCGCTTTTACCCCTTCAGGACGGACCAATACCCATATATATGCGCACGGAAGCTTCT	540	QY	1561	CTTGAGCGTCCGCGCCCTCTCCACATCCAGCATCTCGAAGACCTTCTTTTGCTCTCCG	1620
Db	5627	AGTCCGCTTTTACCCCTTCAGGACGGACCAATACCCATATATATGCGCACGGAAGCTTCT	5686	Db	6707	CTTGAGCGTCCGCGCCCTCTCCACATCCAGCATCTCGAAGACCTTCTTTGCTCTCCG	6766
QY	541	AATATATGCCAGTACCGGGTTGCCGTGCCACAAATCCGTTTACCGCGCTGGTCCCAAT	600	QY	1621	CTCCGGGTAAAGCTCTCTTTCTCGGAGGAGGACACAACTAAAGCCGGTACCTTATAAT	1680
Db	5687	AATATATGCCAGTACCGGGTTGCCGTGCCACAAATCCGTTTACCGCGCTGGTCCCAAT	5746	Db	6767	CTCCGGGTAAAGCTCTCTTTCTCGGAGGAGGACACAACTAAAGCCGGTACCTTATAAT	6826
QY	601	GCTGTGCGCGGTATCGCATCTCATCTCATCTGCGGACAGACCAACACCCCGAAG	660	QY	1681	TATAACACCACTGCTAGCGACCAACTGTTGTGAGAAATGCCCGCGGACACCGGTCGCT	1740
Db	5747	GCTGTGCGCGGTATCGCATCTCATCTCATCTGCGGACAGACCAACACCCCGAAG	5806	Db	6827	TATAACACCACTGCTAGCGACCAACTGTTGTGAGAAATGCCCGCGGACACCGGTCGCT	6886
QY	661	TCCGTTGATGAATTCGAATTAACCTGACGAGTATGCTATGTTAGTCCAGCCGCGATA	720	QY	1741	ATTTCACCTTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTGCGGTT	1800
Db	5807	TCCGTTGATGAATTCGAATTAACCTGACGAGTATGCTATGTTAGTCCAGCCGCGATA	5866	Db	6887	ATTTCACCTTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTGCGGTT	6946
QY	721	GCCTCTGAGCTTGATCCCAAGTGAGCGGCTTACACTATCGTAACCAAGGCTGGCGTCC	780	QY	1801	TTAGCCCCCACCTCTGCGCTAGCATTTGCTTGAGGATACCTTTGACTACCTGCGCGGCC	1860
Db	5867	GCCTCTGAGCTTGATCCCAAGTGAGCGGCTTACACTATCGTAACCAAGGCTGGCGTCC	5926	Db	6947	TTAGCCCCCACCTCTGCGCTAGCATTTGCTTGAGGATACCTTTGACTACCTGCGCGGCC	7006
QY	781	GTGAGACCTCTGGGTGGCTGAGGAGGAGCTTACCTCTGGTCTGTTATGCTTTGCATA	840	QY	1861	CATACCTTTGATGANTTCTGCCAGAGTGC CGCCCTTTGGCCCTTCAGGCTCGGCTTC	1920
Db	5927	GTGAGACCTCTGGGTGGCTGAGGAGGAGCTTACCTCTGGTCTGTTATGCTTTGCATA	5986	Db	7007	CATACCTTTGATGANTTCTGCCAGAGTGC CGCCCTTCAGGCTCGGCTTC	7066
QY	841	CATGGCTCACTCGTAATTCCTACTATATACACCTTATACCGTGGCTCGGGCTGTTG	900	QY	1921	CAGTCTACTGCTCGCTAGCTTACGCGCTTACGATGAAGTGGGTAAAACCTCGGAGTTG	1980
Db	5987	CATGGCTCACTCGTAATTCCTACTATATACACCTTATACCGTGGCTCGGGCTGTTG	6046	Db	7067	CAGTCTACTGCTCGCTAGCTTACGCGCTTACGATGAAGTGGGTAAAACCTCGGAGTTA	7126
QY	901	GACTTTGCCCTTGAGCTTGAGTTTTCGCAACTTACCCCGGTAAACCAATACCGCGGTC	960	QY	1981	TAG 1983	
Db	6047	GACTTTGCCCTTGAGCTTGAGTTTTCGCAACTTACCCCGGTAAACCAATACCGCGGTC	6106	Db	7127	TAG 7129	
QY	961	TCCGTTATTCAGACGCTGCTGCCACCGCTTTCGCGGTGCGGACGGAGCTGCCGAG	1020	RESULT 12			
Db	6107	TCCGTTATTCAGACGCTGCTGCCACCGCTTTCGCGGTGCGGACGGAGCTGCCGAG	6166	HPHEV			
QY	1021	CTCACCACACGCTGCTACCGCTTATGAGGACCTCTATTTTACTAGTACTATGTT	1080	LOCUS	HPHEV	7194 bp	RNA linear VRL 21-MAY-2003
Db	6167	CTCACCACACGCTGCTACCGCTTATGAGGACCTCTATTTTACTAGTACTATGTT	6226	DEFINITION	Hepatitis E virus genomic RNA.		
QY	1081	GTGCTGAGATCGCGCGGGATAGCCCTCACCTGTTCAACCTTCTGTCACACTCTGCTT	1140	ACCESSION	D10330.1	GI:221704	
Db	6227	GTGCTGAGATCGCGCGGGATAGCCCTCACCTGTTTAACTTCTGTCACACTCTGCTT	6286	VERSION	D10330.1	GI:221704	
QY	1141	GGCGGCTCCGACAGAAATGATTTCGTCGCTGGTGGCGAGCTGTTTACTCCGTCCTC	1200	KEYWORDS			
Db	6287	GGCGGCTCCGACAGAAATGATTTCGTCGCTGGTGGCGAGCTGTTTACTCCGTCCTC	6346	SOURCE	Hepatitis E virus		
QY	1201	GTTGCTCAGCAATGGCGAGCGGACTGTTAAGTCTGATACATCTCTAGAGAATGCTCAG	1260	ORGANISM	Hepatitis E virus		
Db	6347	GTTGCTCAGCAATGGCGAGCGGACTGTTAAGTCTGATACATCTCTAGAGAATGCTCAG	6406	REFERENCE	1		
QY	1261	CAGGATAAGGTTATGCAATCCCGCATGACATTTGACCTCGGAGAATCTCGTGGTATT	1320	AUTHORS	Tam, A.W., Smith, M.M., Guerra, M.E., Huang, C.C., Bradley, D.W., Fry, K.E. and Reyes, G.R.		
Db	6407	CAGGATAAGGTTATGCAATCCCGCATGACATTTGACCTCGGAGAATCTCGTGGTATT	6466	TITLE	Hepatitis E virus (HEV) : molecular cloning and sequencing of the full-length viral genome		
QY	1321	CAGGATATGATAACCAACATGAACAGATCGCGACGCTTCTCCAGCCCATCGCGC	1380	JOURNAL	Virology 185 (1), 120-131 (1991)		
Db	6467	CAGGATATGATAACCAACATGAACAGATCGCGACGCTTCTCCAGCCCATCGCGC	6526	MEDLINE	9204067		
QY	1381	CTTTTCTGTCTTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT	1440	PUBMED	1926770		
Db	6527	CTTTTCTGTCTTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT	6586	REFERENCE	2 (bases 1 to 7194)		
				AUTHORS	Aye, T.T., Uchida, T., Ma, X., Iida, F., Shikata, T., Ichikawa, M., Rikihisa, T. and Win, K.M.		
				TITLE	Sequence and gene structure of the hepatitis E virus isolated from Myanmar		
				JOURNAL	Virus Genes 7 (1), 95-109 (1993)		
				MEDLINE	93227573		
				PUBMED	8470371		
				REFERENCE	3 (bases 1 to 7194)		
				AUTHORS	Uchida, T.		
				TITLE	Direct Submission		
				JOURNAL	Submitted (23-JAN-1992) Toshikazu Uchida, Nihon University School of Medicine, 1st Department of Pathology; 30-1 Oyaguchi-Kaminachi,		

Itabashi-Ku, Tokyo 173, Japan (Tel:81-3-3972-8111,
Fax:81-3-3972-8830)

Submitted (23-Jan-1992) to DDBJ by:

Toshikazu Uchida

Ist Department of Pathology

Nihon University School of Medicine

30-1 Oyaguchi Kamimachi

Itabashi-Ku, Tokyo 173

Japan

Phone: 81-3-3972-8111

Fax: 81-3-3972-8830

Location/Qualifiers

1. .7194

/organism="Hepatitis B virus"

/mol_type="genomic RNA"

/isolate="HEVNE8L"

/db_xref="taxon:12461"

28. .5109

/codon_start=1

/product="ORF1"

/protein_id="BAA01172.1"

/db_xref="GI:221705"

/translation="MEAHQFIKAHGITTAEQALAAANSALANAVVRPFUSHOOIE
ILINMQPQLVREVEFWNHPIQRIHNELELYCRASGRCLGTGAHPRSDNPNV
VHRCFLPVGRDQVWYAPTAPRPAACRRSALRGLPAVDRTYCLDGFSGCNFPAETG
IALYSLHDSMPSDVAEAMFRHGMRLYAAHLPLPEVLPPTGYRTASVLLIHDGRV
VTEGDTAGYNHNSLRWIRTTKVTDHPLVIERVRAIGCHEFVLLITAAPEPSM
PYVPRSTEVYVRSIFGPGTSPISFPTSCSTKSTFAHVAHIDRLMLFGATLDDQA
FCCSRMLYRGISVYVGLTVANEGWNASDALAVITAAIYLTICHOYLRTQAS
KGMRLREHQAFTITPYSWLFKSGRDYIEGRQLEFYAQRRLWSAGFHLDPRVL
FDESAPCRCTAIRKALQKFCFMKWLGOECTFLQPAEGVGVGDQGHNEAYEGSDVD
PAESAISDTSYVVPGLTQALDPLPAIIVARAGRLTATVKSVDGIDCELT
LGNKTRFSPVDGVALEANGPERYNSLFDASOSTMAAGPFSLTAAAGLEVRVAA
GLDHRAVAPGVSPEGVEVAPCSALYRFRNBAQRSLTGNLWPHBGLLGLFAPP
SPGHVESANPGSSTLYRTWSEVDVSSPARPDGLMSEPSIPSRRAATPLAVLL
PPAPDPDPSPALDEPASPAGATAGAPAITHTQARHRLLLFTYDPSKVFAGSLFES
TCTVLVNASVNDHRRPGGLCHAFYQRYPASFDASFMVRDGAAYTLTPRPIHVA
DYRLHNPKEARLEAAYRETCRLGTAAYSLLGTGIYQVPIGPSFDWERNRHPGDELY
PELAARMEANPRTPLTITEDVARTANLAIELSDATVGBACARVTPGVQYOF
TAGVPGSKRSITQADVUVVUPPRELNRNRGFAAFNPHTHAARTVQGRVIDE
APSLPHLLHLMQATVHLLGNQIPADIDFEHAGLVPAIRPDLGPTSMWHVTHRC
PADVCELIRGAPMTQTSRVLRSIFWGPAPVQKLVFTQAAKANPSGVTVHEAQA
TYETITATDARGILQSSRAHAIVALTRHTEKFIIDAPGLLREVIGSDAIIVNFF
LAGGELHGORPASI PRGNDAVDTLAAPPSCQISAPHOLAEELGHRVVPVAAVLPP
CEPLBOGLLYLPGHATCDVSUTFEITDILVHCMAAPNOKKAVLSTLVGRYGRKLY
NASHSDRSLARFTIAGPVOVTTCELYELVEAMVEKQDGSVLELIDLNRDVSRI
TFQXQCNKFTGETIAGKVGQGISAMSKTFCALFGPMFRAIEKAILALLPQGVYF
DAFDDTVSAAVAAKASVNFENFSEFSTQNNFSLGLECAIMEECQPOWLIRLYH
LIRSAWILQAPKESLRGFWKHSGEPGTLIMNTVNMMAVITHCYDFRDPQVAAFQDD
SIVLCSYRQSPGAALVLAGCLLKVDPRPIGLVAGVVVAPGLGALPDVVPFAGRLT
EKNMGPGPERASQBLAVSDFLRLKLTNVAQMCVDVSVRYGVSPGLVHNLIGMLQAVA
DGKAFTSVKVLDTNLSILCRVE"

5106. .5477

/codon_start=1

/product="ORF3"

/protein_id="BAA01173.1"

/db_xref="GI:221706"

/translation="MNNMSFAAPMGSRPCDLGLFCCSCFCCLCPHRHPVSRLLAAV
GGAAVPAVSGVTGLIUSPSQSPIFIQTPSPMSPLEPGLDLVFANQPDHSAPLGV
TRPSAPPLPVHVDLPLQGRPR"

5147. .7129

/codon_start=1

/product="ORF2"

/protein_id="BAA01174.1"

/db_xref="GI:221707"

/translation="MKRPRPIILLMLPAPPQSPGRRRRSGSGGGFWMD
RVDSOPFAPYTHPTNPPAPVTAAGAPRQAPRPLGSAAWDAQORPAVARSRRP
TTAGAPLTAVAHADPTPVPDVSFAAILRQYNLSTPLTSSVATGNLVLVYAAPL
SELLPLQDGTNTHIMATESYAOQVVRATIRYRPLVNAVGGYIAISISFWPQTITI
PVSVMNSIISDVRLLVQGIASELVI FSELRHYNQGRWSVETSGVABEEATSLGV
MLCIHGPSVNSYNTPTGALGLDPALEFLNLTGNTNTRVSRYSSTARRHLRG
ADGTALTTAATRFMKDLYFTSTNGVIGIRGIALTLFNLADTLGLGPTLISAG
GQLFYSRPVSAHGEPTVKLYTSVENAQDKGIAIPHDIDLGESRVVIQDYNQHEQD

RPTSPAPSPRPFSLVRANDVLWLSLTAABYDQSTYSSSTAPVYVSDSVTLNVNATGAQ
AVARSLDWTKVTLDGRPLSTTQQVPTFFVLPRGLKLSFWEAGTTKAGYPYNYNTAS
DQLLVENAAHRAVLTSTVTSILGAGPVSISAVAVLPHASALLLEDTLTDYPACATFD
DFCEPCRLPLGQCAQFSTVAELQRLKMKVGTREL"

7194

polyA_site

ORIGIN

Query Match 97.6%; Score 1935; DB 14; Length 7194;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 1953; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY	1	ATGCGCCCTCGGCCCTATTTTGTGTCTCATGTCTTTTTCCTATGCTGCCCGGCCA	60
DB	5147	ATGCGACCTCGGCCCTATTTTGTGTCTCATGTCTTTTTCCTATGCTGCCCGGCCA	5206
QY	61	CGCGCCGCTCAGCCGCTGCTGGCGCGCGTCTGGCGCGCGAGCGGCTTCCCGCGGTGT	120
DB	5207	CGCGCCGCTCAGCCGCTGCTGGCGCGCGTCTGGCGCGCGAGCGGCTTCCCGCGGTGT	5266
QY	121	TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC	180
DB	5267	TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC	5326
QY	181	CCCTTTCGCCCCCGATGTCAACGCTCGCGCGCGGCTGACCTCTGCTTTCGCAACCCG	240
DB	5327	CCCTTTCGCCCCCGATGTCAACGCTCGCGCGCGGCTGACCTCTGCTTTCGCAACCCG	5386
QY	241	CGAACCTCTCGGCTCGGCTTGGCGTCAACGAGCGCGAGCGCGCGGCTTTCGCTTTC	300
DB	5387	CGAACCTCTCGGCTCGGCTTGGCGTCAACGAGCGCGAGCGCGCGGCTTTCGCTTTC	5446
QY	301	AGACCTTACACAGCTGGGGCGCGCGCTTAAACCGCGCTCGCTCCGCGCCCATGACAC	360
DB	5447	AGACCTTACACAGCTGGGGCGCGCGCTTAAACCGCGCTCGCTCCGCGCCCATGACAC	5506
QY	361	CCAGTGCCTGTGTGTGACTTCCCGCGCGCCATCTTTCGCGCGGAGTATTAACCTATCA	420
DB	5507	CCAGTGCCTGTGTGTGACTTCCCGCGCGCCATCTTTCGCGCGGAGTATTAACCTATCA	5566
QY	421	TTCTCCCTTACCTCTTCGCTGGCGCACCGGCACCTAACCTGGTCTTTTATCGCGCCCT	480
DB	5567	TTCTCCCTTACCTCTTCGCTGGCGCACCGGCACCTAACCTGGTCTTTTATCGCGCCCT	5626
QY	481	AGTTCGCTTTTACCCCTTCAGGACGCGACCAATACCATATAATGGCCACGGAAGTCT	540
DB	5627	AGTTCGCTTTTACCCCTTCAGGACGCGACCAATACCATATAATGGCCACGGAAGTCT	5686
QY	541	AATTATGCCAGTACCGGGTTGCGCGTGCACCAATTCGCTTACCGCGCGTGTGCCCAAT	600
DB	5687	AATTATGCCAGTACCGGGTTGCGCGTGCACCAATTCGCTTACCGCGCGTGTGCCCAAT	5746
QY	601	GCTGTGGGGGTTACGCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCGCGAG	660
DB	5747	GCTGTGGGGGTTACGCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCGCGAG	5806
QY	661	TCGCTGTATGATGAATTCATTAACCTTCGACGATGTTTGTATTTAGTCAGCGCCGCA	720
DB	5807	TCGCTGTATGATGAATTCATTAACCTTCGACGATGTTTGTATTTAGTCAGCGCCGCA	5866
QY	721	GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTTACACTATCGTAAACCAAGGCTGG	780
DB	5867	GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTTACACTATCGTAAACCAAGGCTGG	5926
QY	781	GTCGAGACCTCTGGGGTGTGCTGAGGAGGCTACTCTGGTCTTGTATGCTTTCGCA	840
DB	5927	GTCGAGACCTCTGGGGTGTGCTGAGGAGGCTACTCTGGTCTTGTATGCTTTCGCA	5986
QY	841	CATGGCTCACTCGTAAATTCCTATTAACACCTATACCGGTCCTCGGCGCTGTG	900
DB	5987	CATGGCTCACTCGTAAATTCCTATTAACACCTATACCGGTCCTCGGCGCTGTG	6046
QY	901	GACTTTCGCTTGTGAGTTGAGTTTCGCAACCTTACCCCGGTTAACCAACATACGCGG	960

```
Db 6047 GACTTTGCCCTTGAGCTTGAGTTTGGCAACTTACCCCGGTAAACCAATAACGCGGTC 6106
Qy 961 TCCGGTATTTCAGCACTGCTCGCCACCGGCTTGTGCGGTGCGGAGCGGACTGCCAG 1020
Db 6107 TCCGGTATTTCAGCACTGCTCGCCACCGGCTTGTGCGGTGCGGAGCGGACTGCCAG 6166
Qy 1021 CTCACCAACCAAGGCTGCTACCGCTTATGAAGGACCTCTATTATCTAGTACTAATGGT 1080
Db 6167 CTCACCAACCAAGGCTGCTACCGCTTATGAAGGACCTCTATTATCTAGTACTAATGGT 6226
Qy 1081 GTGCGTGAATGCGCGCGGATAGCCCTCACCCCTGTTCACCTGTGCTGACCTGCTGCTT 1140
Db 6227 GTGCGTGAATGCGCGCGGATAGCCCTCACCCCTGTTCACCTGTGCTGACCTGCTGCTT 6286
Qy 1141 GCGGCTGCGGACAGAAATTGATTTCTGCGCTGGTGGCCAGCTGTTCTACTCCCGTCCC 1200
Db 6287 GCGGCTGCGGACAGAAATTGATTTCTGCGCTGGTGGCCAGCTGTTCTACTCCCGTCCC 6346
Qy 1201 GTTGTCTCAGCCCAATGCGGAGCGGACCTGTTAAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db 6347 GTTGTCTCAGCCCAATGCGGAGCGGACCTGTTAAAGTTGTATACATCTGTAGAGAATGCTCAG 6406
Qy 1261 CAGATAAAGGATATTGCAATCCCGCATGACATGACCTCGGAGAAATCTCGTGTGTTATT 1320
Db 6407 CAGATAAAGGATATTGCAATCCCGCATGACATGACCTCGGAGAAATCTCGTGTGTTATT 6466
Qy 1321 CAGGATTATGATTAACCAACATGACACAGATCGCGGAGCGCTTCTCAGACCCCATCGCGC 1380
Db 6467 CAGGATTATGACCAACCAACATGACACAGATCGCGGAGCGCTTCTCAGACCCCATCGCGC 6526
Qy 1381 CCTTTCCTCTGCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
Db 6527 CCTTTCCTCTGCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 6586
Qy 1441 GACAGTCCACTATATGCTCTTCGACTGCGGCCAGTTTATGTTCTGACTCTGTGACCTTG 1500
Db 6587 GACAGTCCACTATATGCTCTTCGACTGCGGCCAGTTTATGTTCTGACTCTGTGACCTTG 6646
Qy 1501 GTTAAATGTCGACCGCGCGGAGCGGCTGCGGCTGCGCTCCGATTCGACCAAGGTACA 1560
Db 6647 GTTAAATGTCGACCGCGCGGAGCGGCTGCGGCTGCGCTCCGATTCGACCAAGGTACA 6706
Qy 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCCTGCGG 1620
Db 6707 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCCTGCGG 6766
Qy 1621 CTCGCGGTAAGCTCTCTTTCTGGGAGCGAGGACAACTAAAGCGGGTACCTTATAAT 1680
Db 6767 CTCGCGGTAAGCTCTCTTTCTGGGAGCGAGGACAACTAAAGCGGGTACCTTATAAT 6826
Qy 1681 TATAACACCACTGCTAGCGACCAACTGCTTTGTGAGAAATCCGCGCGGACCGGGTCCGT 1740
Db 6827 TATAACACCACTGCTAGCGACCAACTGCTTTGTGAGAAATCCGCGCGGACCGGGTCCGT 6886
Qy 1741 ATTTTCCACTTACACCACTAGCCTGGGTGCTGGTCCGCTCTCCATTTCTCGGGTTGCGGTT 1800
Db 6887 ATTTTCCACTTACACCACTAGCCTGGGTGCTGGTCCGCTCTCCATCTCTGCGGTGCTGTT 6946
Qy 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGATACCTTGAGATACCTGCGCGCGCC 1860
Db 6947 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGATACCTTGAGATACCTGCGCGCGCC 7006
Qy 1861 CATACTTTTTCATGATTTCTGCCAGAGTGCAGCGCCCTGCGCTTCAGGGCTCGCTTTC 1920
Db 7007 CATACTTTTTCATGATTTCTGCCAGAGTGCAGCGCCCTTTCAGGGCTCGCTTTC 7066
Qy 1921 CAGTCTACTGCTGAGCTTTCAGCGCTTTCAGTGAAGTGGGTGAGTAACTCGGGAGTTG 1980
Db 7067 CAGTCTACTGCTGAGCTTTCAGCGCTTTCAGTGAAGTGGGTGAGTAACTCGGGAGTTA 7126
Qy 1981 TAG 1983
Db 7127 TAG 7129
```

RESULT 13

```
AF124407
LOCUS      1983 bp      RNA      linear      VRL 24-JUL-1999
DEFINITION Hepatitis E virus isolate AKL-90 ORF-2 gene, complete cds.
ACCESSION  AF124407
VERSION     AF124407.1  GI:5579229
KEYWORDS
SOURCE
  ORGANISM  Hepatitis E virus
            Hepatitis E virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Hepatitis
            E-like viruses.
REFERENCE   1 (bases 1 to 1983)
AUTHORS    Arankalle,V.A., Paranjape,S., Emerson,S.U., Purcell,R.H. and
            Walimbe,A.M.
TITLE      Phylogenetic analysis of hepatitis E virus isolates from India
            (1976-1993)
JOURNAL    J. Gen. Virol. 80 (Pt 7), 1691-1700 (1999)
MEDLINE    9350000
PUBMED     10423137
REFERENCE   2 (bases 1 to 1983)
AUTHORS    Arankalle,V.A., Paranjape,S., Emerson,S.U., Purcell,R.H. and
            Walimbe,A.M.
TITLE      Direct Submission
JOURNAL    Submitted (28-JAN-1999) Hepatitis, National Institute of Virology,
            20-A Dr. Ambedkar Road, Pune, Maharashtra 411001, India
FEATURES
  source
    1..1983
        /organism="Hepatitis E virus"
        /mol_type="genomic RNA"
        /isolate="AKL-90"
        /db_xref="taxon:12461"
        /country="India"
    1..1983
        /codon_start=1
        /product="ORF-2"
        /protein_id="AAD45493.1"
        /db_xref="GI:5579230"
        /translation="MRPRPILLLLMFLPMLPAPPQPQSRRRRRSGSGGGPFGD
            RVDQFALPYIHPNTPFADVTAAAGARRINQAPRLGSAWRDQAPPAASRRRP
            TTAGAALPLAVAPADTPVDVDSGAILRRQYNLSPLTSSVATGNLVLPAFL
            SPLPLQDGTNTHIMATASNYARVATIRPLVPNAVGSYISFWPQTTLV
            PTVDMNSITTDVILVQPIASELVIPIERLHYRNOQWRSVETSGVABEATGLV
            MLCIGSPNSVNTPTYGALGLDLEPENTPGNNTVSVSVSTARHLRSG
            AGTARLTATTAATREMKDLYFTSTNGVIGRGIALTLNLATLLGLPTELISSGQ
            GQLFYRPPVWSANGEPTVKLYTSVENAQDDKGIAPHDIDLGESRVVIQYDNOHQD
            RPTSPAPRPFSLRANDVLMLSLTAAYDQSTGSGTSPVTVSDSVTLNVATGAQ
            AVARSLDWTKVTLDGRLSTIQOYSKTFVLPLRKLFWAGTTCAGTTCAGTTCAGT
            DOLLVNAAGHRVAISTYTTSLGAGPVSISAVAVAFAPHSALALLEDTLDYPARAHTFD
            DFCPECPRLGQCAFQSVAEIQLRLKMKVGKTRFL"
```

ORIGIN

```
Query Match      96.5%; Score 1914.2; DB 14; Length 1983;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1940; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 ATGCGGCTCGGCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCGCGGCCA 60
Db 1 ATGCGGCTCGGCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCGCGGCCA 60
Qy 61 CGCGCCGCTCAGCGCTGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCTTCCGCGGTTG 120
Db 61 CGCGCCGCTCAGCGCTGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCTTCCGCGGTTG 120
Qy 121 TTCTGGGGTGACCGGTTGATTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 121 TTCTGGGGTGACCGGTTGATTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Qy 181 CCCTTCGCCCCCGATGTCACCGCTCGCGCGCGGCTGGACCTGCTGTCGACACCGCC 240
Db 181 CCCTTCGCCCCCGATGTCACCGCTCGCGCGCGGCTGGACCTGCTGTCGACACCGCC 240
```

241 CGACCACTCGGCTCCGCTTGGCTGACACGAGCCCGAGCCCGCGCTTGCCTCAAGTGGT 300
Db |||||
241 CGACCACTCGGCTCCGCTTGGCTGACACGAGCCCGAGCCCGCGCTGCTCAAGTGGT 300
Qy |||||
301 AGACCTACACAGCTGGGGCGCGCGGTAAACCGGGTGGCTCGCGGCCATGACACCCGG 360
Db |||||
301 AGACCTACACAGCTGGGGCGCGCGGTAAACCGGGTGGCTCGCGGCCATGACACCCGG 360
Qy |||||
361 CCAGTGCTGATGTCGACTCCGCGGGCGGCATCTTTGGCGGCGAGTATACCTATCAACA 420
Db |||||
361 CCAGTGCTGATGTCGACTCCGCGGGCGGCATCTTTGGCGGCGAGTATACCTATCAACA 420
Qy |||||
421 TCTCCCTTACCTCTTCCTGGTGGCCACCGGCACTAACCTGGTCTTTATGCCGCCCTCTT 480
Db |||||
421 TCTCCCTTACCTCTTCCTGGTGGCCACCGGCACTAACCTGGTCTTTATGCCGCCCTCTT 480
Qy |||||
481 AGTCCGCTTTTACCCTTTCAGAGCGGCAACATACCTATATATGTCGACGGAAGTCTT 540
Db |||||
481 AGTCCGCTTTTACCCTTTCAGAGCGGCAACATACCTATATATGTCGACGGAAGTCTT 540
Qy |||||
541 AATTATGCCAGTACCGGGTTCGCGGTGCGCAATCCGTTACGCGCGGTGCTCCCAAT 600
Db |||||
541 AATTATGCCAGTACCGGGTTCGCGGTGCGCAATCCGTTACGCGCGGTGCTCCCAAT 600
Qy |||||
601 GCTGTCGGCGGTACGCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCCGGAG 660
Db |||||
601 GCTGTCGGCGGTACGCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCCGGAG 660
Qy |||||
661 TCCGTTGATATGAATTCATTAACCTCGACGGATGTTGTTATTTAGTCAGCCCGGCATA 720
Db |||||
661 TCCGTTGATATGAATTCATTAACCTCGACGGATGTTGTTATTTAGTCAGCCCGGCATA 720
Qy |||||
721 GCCTCGAGCTGTGATCCCAAGTGAAGCCCTACACTATCGTAACCAAGGCTGCGCTCC 780
Db |||||
721 GCCTCGAGCTGTGATCCCAAGTGAAGCCCTACACTATCGTAACCAAGGCTGCGCTCC 780
Qy |||||
781 GTCGAGACCTCGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
Db |||||
781 GTCGAGACCTCGGGGTGGCTGAGGAGGAGGCACTCTCGGTCTTGTATGCTTTGCATA 840
Qy |||||
841 CATGGCTACTCGTAAATTCCTATATAACCTATACCGGTGCTCGGCTGGTGGT 900
Db |||||
841 CATGGCTACCCGTAATTCCTATATAACCTATACCGGTGCTCGGCTGGTGGTGGT 900
Qy |||||
901 GACTTTGGCCTTGAAGTTTGGAGTTTGGCAACCTTACCCCGGTAACCAATACGGGGTC 960
Db |||||
901 GACTTTGGCCTTGAAGTTTGGAGTTTGGCAACCTTACCCCGGTAACCAATACGGGGTC 960
Qy |||||
961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCTGTCGGTGGCGGAGCGGAGTGCAG 1020
Db |||||
961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCTGTCGGTGGCGGAGCGGAGTGCAG 1020
Qy |||||
1021 CTCACACCAACCGCTGCTACCGCTTTATGAAGGAGCTCTATTTACTAGTACTAATGGT 1080
Db |||||
1021 CTCACACCAACCGCTGCTACCGCTTTATGAAGGAGCTCTATTTACTAGTACTAATGGT 1080
Qy |||||
1081 GTCGGTGAATCGCGGGGATAGCCCTACCCCTGTTCAACCTGCTGACACTCTGCTT 1140
Db |||||
1081 GTCGGTGAATCGCGGGGATAGCCCTACCCCTGTTCAACCTGCTGACACTCTGCTT 1140
Qy |||||
1141 GCGCGCTTCCGACAGAAATTTGTTGTCGGTGGTGGCGAGCTGTTCTACTCCCGTCCC 1200
Db |||||
1141 GCGCGCTTCCGACAGAAATTTGTTGTCGGTGGTGGCGAGCTGTTCTACTCCCGTCCC 1200
Qy |||||
1201 GTTGTCTACGCAATGGCGAGCGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
Db |||||
1201 GTTGTCTACGCAATGGCGAGCGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
Qy |||||
1261 CAGGATAAGGTTATTCATTCGCGATGACATGACCTCGGAAATCTCGTGGTGTATT 1320
Db |||||
1261 CAGGATAAGGTTATTCATTCGCGATGACATGACCTCGGAAATCTCGTGGTGTATT 1320
Qy |||||
1321 CAGGATTATGATAACCAATGAACAGATCGCGCGAGCGCTTCTCCAGCCCATCGCGC 1380

Db |||||
1321 CAGGATTATGATAACCAATGAACAGATCGCGCGAGCGCTTCTCCAGCCCATCGCGC 1380
Qy |||||
1381 CTTTCTCTGCTCTGAGCTAAATGATGCTTTGGCTCTCTCTCACCCTGCGAGTAT 1440
Db |||||
1381 CTTTCTCTGCTCTGAGCTAAATGATGCTTTGGCTCTCTCTCACCCTGCGAGTAT 1440
Qy |||||
1441 GACCACTGCTTATGGCTCTTCGACTGCGCCAGTTTATGTTCTGACTCTGACTG 1500
Db |||||
1441 GACCACTGCTTATGGCTCTTCGACTGCGCCAGTTTATGTTCTGACTCTGACTG 1500
Qy |||||
1501 GTTAAATGTTGGACGCGCGCGAGCTGTCGCCGTGCTCGATTTGACCAAGGTACA 1560
Db |||||
1501 GTTAAATGTTGGACGCGCGCGAGCTGTCGCCGTGCTCGATTTGACCAAGGTACA 1560
Qy |||||
1561 CTTGAGGCTGCGCCCTCTCCACATCCAGAGTACTCGAAGACCTTCTTTGCTCGCG 1620
Db |||||
1561 CTTGAGGCTGCGCCCTCTCCACATCCAGAGTACTCGAAGACCTTCTTTGCTCGCG 1620
Qy |||||
1621 CTTCCGCGTAACTCTCTTTCTGGAGGAGGAGCAAACTAAAGCCGGGTACCTTATAAT 1680
Db |||||
1621 CTTCCGCGTAACTCTCTTTCTGGAGGAGGAGCAAACTAAAGCCGGGTACCTTATAAT 1680
Qy |||||
1681 TATAACACCACTGCTAGCGACCACTGCTTGTGAGAAATGCGCGGCGACCGGTCGT 1740
Db |||||
1681 TATAACACCACTGCTAGCGACCACTGCTTGTGAGAAATGCGCGGCGACCGGTCGT 1740
Qy |||||
1741 ATTTCACCTTACACCACTAGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1800
Db |||||
1741 ATTTCACCTTACACCACTAGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1800
Qy |||||
1801 TTAGCCCCCACTCTGCGCTAGCATTTGTTGAGGATACCTTGGACTACCTGCCCGCGC 1860
Db |||||
1801 TTAGCCCCCACTCTGCGCTAGCATTTGTTGAGGATACCTTGGATACCTGCCCGCGC 1860
Qy |||||
1861 CATACTTTTGAATTTCTGCCAGAGTCCGCCCTTGGCTTTCAGGGCTGCGCTTTC 1920
Db |||||
1861 CATACTTTTGAATTTCTGCCAGAGTCCGCCCTTGGCTTTCAGGGCTGCGCTTTC 1920
Qy |||||
1921 CAGTCTACTGCTGCTGAGCTTACAGCCCTTAAAGTGAAGGTGGGTAAACTCGGAGTTG 1980
Db |||||
1921 CAGTCTACTGCTGCTGAGCTTACAGCCCTTAAAGTGAAGGTGGGTAAACTCGGAGTTG 1980
Qy |||||
1981 TAG 1983
Db |||||
1981 TAG 1983

RESULT 14
AF051830

LOCUS AF051830 7199 bp RNA linear VRL 18-DEC-1998
DEFINITION Hepatitis E virus strain TK15/92, complete genome.

ACCESSION AF051830

VERSION AF051830.1 GI:4033730

KEYWORDS

SOURCE Hepatitis E virus

ORGANISM Hepatitis E virus

Viruses; ssRNA positive-strand viruses, no DNA stage; Hepatitis E-like viruses.

REFERENCE 1 (bases 1 to 7199)

AUTHORS Gouvea V., Snellings, N., Popek, M.J., Longer, C.F. and Innis, B.L.

TITLE Hepatitis E virus: complete genome sequence and phylogenetic

analysis of a Nepali isolate

JOURNAL Virus Res. 57 (1), 21-26 (1998)

MEDLINE 99049628

PUBMED 9833882

REFERENCE 2 (bases 1 to 7199)

AUTHORS Gouvea V.

TITLE Direct Submission

JOURNAL Submitted (02-MAR-1998) Department of Virus Diseases, Walter Reed

Army Institute of Research, 14th and Dahlia St., Bldg. 40,

Washington DC 20307-5100, USA

FEATURES Location/Qualifiers

source	1. .7199	Query Match	95.6%;	Score 1896.6;	DB 14;	Length 7199;
	/organism="Hepatitis E virus"	Best Local Similarity	97.3%;	Pred. No. 0;		
	/mol_type="genomic RNA"	Matches 1929;	Conservative 0;	Mismatches 54;	Indels 0;	Gaps 0;
	/strain="TK15/92"					
	/db_xref="taxon:12461"					
	/country="Nepal"					
	/note="isolated from the bile of experimentally infected cynomolgous monkey"					
CDS	28. .5109	Qy	1	ATCGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTGTGCTATGCTATGCTGCGCGGCCA	60	
	/notes="encodes cysteine protease, RNA helicase and RNA-dependent RNA polymerase; ORF1"	Db	5147	ATCGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTGTGCTATGCTATGCTGCGCGGCCA	5206	
	/codon_start=1	Qy	61	CCGCGCGCTCAGCCGCTTGGCCCGCGCTCGTGGCGGGGCGAGCGCGGTTCGGCGGGTGGT	120	
	/product="nonstructural polyprotein"	Db	5207	CCGCGCGCTCAGCCGCTTGGCCCGCGCTCGTGGCGGGGCGAGCGCGGTTCGGCGGGTGGT	5266	
	/protein_id="AAC97186.1"	Qy	121	TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTGGCAATCCCTTATATTCATCCAAACCAAC	180	
	/db_xref="GI:4033731"	Db	5267	TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTGGCAATCCCTTATATTCATCCAAACCAAC	5326	
	/translation="MEAHQFIKAPGITTAIEQALAAANSALANAVVVRPFLSHQOIE	Qy	181	CCCTTCGCCCCCGATGTCAACCGCTGCGCGCGGGCTGGACCTCGTGTGCGCCAAACCGGCC	240	
	ILINLMPQVLPFRPVRVNHPIQRIVHNELELYCRASGRCLIEI GAHPRSSINDNPNV	Db	5327	CCCTTCGCCCCCGATGTCAACCGCTGCGCGCGGGCTGGACCTCGTGTGCGCCAAACCGGCC	5386	
	VHCFERLPQVDRVRYTAPTGRPAANCRSRALRGLPAADRTYCFDGFSGSCSPSTGV	Qy	241	CGACCACTCGGCTCCCGCTTGGCGTGGACCGGCCAGCGCGCGCGCTGGCTTGCCTCACGTCGT	300	
	IALLYSLHDSNPSDAVAMFRHGMTRLYAALHLPPEVLLPGTYRTASYLIIHIDGRVY	Db	5387	CGACCACTCGGCTCCCGCTTGGCGTGGACCGGCCAGCGCGCGCGCTGGCTTGCCTCACGTCGT	5446	
	VYEGTSGAGYNHDSNLSRWIRTKVTGDHPLVIERVRAIGHCFHLLLTAAPEPSPM	Qy	301	AGACCTTACCACAGCTGGGCGCGCGCGCTAACCGCGGTGCTCCGGGCCCATGACACCCCGG	360	
	PVYPVPRSTEVVRSIFGPGGTSLPSTSCSTKSTFHAVPAHIWDRIMLQFGLTDQQA	Db	5447	AGACCTTACCACAGCTGGGCGCGCGCGCTAACCGCGGTGCTCCGGGCCCATGACACCCCGG	5506	
	FCCSRMLTYLGI SYKVTGTLVANGWRASEDALTAIVITAAALITICHQRYLRTQALS	Qy	361	CCAGTGCCTGATGTGCACTCCCGCGCGCGCATCTTGGCGCGGAGTATACCTATCAACA	420	
	KGNRLRERHAQFIITRLYSWLFPEKSGRDIYIPQRIEYFAQCRRLWSAGFHLDPRLVY	Db	5507	CCGGTGCCTGATGTGCGCTCCCGCGCGCGCATCTTGTGCGCGGAGTATACCTATCAACA	5566	
	FDASAFCHRTVIRKLSKFCPCFMKLGQECTFLQPAEYVGGDGHNEAYEGSDVD	Qy	421	TCCTCCCTTACCTCTTCGCGTGGGCAACCGGCTAACCTGCTGCTTTTATGCGCGCCCTCTT	480	
	PAESAIISDGSVVVPGTALQIYQALDLPABEIVARACRLTATVKYQSOADGRIDCETL	Db	5567	TCCTCCCTTACCTCTTCGCGTGGGCAACCGGCTAACCTGCTGCTTTTATGCGCGCCCTCTT	5626	
	LGKTRTSPVDCAVLETNGRHNLSFQASQSTMAAGPFLTYAASAGLEVRYVAA	Qy	481	AGTCCGCTTTTACCCCTTCAGAGCGGCAACCAATACCATATATATGCGCCAGGAGCTTCT	540	
	GLDHRAIFAPGVSPSTPGVTAFCSALVFENAAORHSLTGNLWFHEPLGLFAPF	Db	5627	AGTCCGCTTTTACCCCTTCAGAGCGGCAACCAATACCATATATATGCGCCAGGAGCTTCT	5686	
	SPGHVESANPFGESTLYTRTWSEIDAVSSPAPDLGVSEPSIESRAATPFLAAPL	Qy	541	AATATGCGCGAGTACCGGGTGGCGGTGCGCAACCTCGGTTACCGCGCGCTGGTCCCAAT	600	
	PPAPDPSPPPADPEPAGPTAGAPALITHQMARHRLLLFTYDGSKVFAAGLSPES	Db	5687	AATATGCGCGAGTACCGGGTGGCGGTGCGCAACCTCGGTTACCGCGCGCTGGTCCCAAT	5746	
	TCWLVANVDRHPPGGGLHAFYFPASFDAAAFVMDGAAAYTLTPRPIHAPV	Qy	601	GCCTGCGCGGGTTACGCCATCTCCATCTCATTTCTGCGCCACAGACCAACCCCGGAGC	660	
	DYLEHNKPLEAARETCSRGTAAAYPLIGTIYQVPIGPSFDAMERNHRPGDELYL	Db	5747	GCCTGCGCGGGTTACGCCATCTCCATCTCATTTCTGCGCCACAGACCAACCCCGGAGC	5806	
	PELAARFANRPTPLIITEDVANTANLAIELDSATDVGRACAGCTVPGVVOYQF	Qy	661	TCCGTTGATATGAATTCATTAACCTCGACGGATGTTTCTGATTTTAGTCCAGCCCGGAGT	720	
	TAGVPGSGKRSITQADVDVVVPTRELNAWRERGFAAFTHTAATVGLPAIRPDLAPSWHVRHC	Db	5807	TCCGTTGATATGAATTCATTAACCTCGACGGATGTTTCTGATTTTAGTCCAGCCCGGAGT	5866	
	APSLPHILLNMQIIPALIDEPHAGLVPALIRPDLAPSWHVRHC	Qy	721	GCCTCTGAGCTTGTGATCCCAAGTACGCGCTACACTTCGTAAACCAAGGCTGGCGCTCC	780	
	PADVCILLIYPMIQTTSRLSLFWGEPVQKLVTPQAGKAAPGVSIVTVEHQA	Db	5867	GCCTCTGAGCTTGTGATCCCAAGTACGCGCTACACTTCGTAAACCAAGGCTGGCGCTCC	5926	
	LYTETITITADARGITQSSRAHAI VALTRHTEKCVI I DAPGLLEVRGSI DSAIVNFF	Qy	781	GTGAGACCTCTCGGGTGGCTGAGGAGGAGCTACCTCTGCTTCTGTTATGCTTTTGATA	840	
	CPLBEOGLLYLPOELTCDSTVFETLTVHCRMAAPSORKAVLSITLVGRYGRRTKLY	Db	5927	GTGAGACCTCTCGGGTGGCTGAGGAGGAGCTACCTCTGCTTCTGTTATGCTTTTGATA	5986	
	NASHDRDLSARFI PAIGPVQVTTCELCHEVMEKQDGSALVELDLGNDRDVSRI	Qy	841	CATGGCTCTCACTCGTAAATTCCTATCTACTATAACACCTTATACCGGTGCCCTCGGGCTGTTG	900	
	TFPKQCNKTGCTETLVHGVKGGIISAMSKTFCALPGMPRAIEKAILLPOGVSYG	Db	5987	CATGGCTCTCACTCGTAAATTCCTATCTACTATAACACCTTATACCGGTGCCCTCGGGCTGTTG	6046	
	DARDDTVFSAVAAKASVFNDFSEFSTQNNFTNSLGLCEAMCGGQMPQWLIRLYH	Qy	901	GACTTTGCGCTTGAGCTTGAGTTTCCGACCTTACCCCGGCTAACCAATACCAATACCGGGTTC	960	
	LIRSAITLAPKESLRGFKKHSGEFTLLTNVNNNAVITHCYDFRDLQVAAFPQDD	Db	6047	GACTTTGCGCTTGAGCTTGAGTTTCCGACCTTACCCCGGCTAACCAATACCAATACCGGGTTC	6106	
	SIVLCSEYQSGAAVLIACGGLKLVDFRPIGLYAGVYVAPGLGALPDVVRPAGRLT	Qy	961	TCCCGTATTTCACAGCACTGCTCGCCACCGCTTCGTCGCGGTGGGAGCGGAGTCCGCGAG	1020	
	EKNWGPPEAEQRLAVSDFLRKLTNAQCMVDVSVRVYVSGPLVHNLIGLQVA	Db	6107	TCCCGTATTTCACAGCACTGCTCGCCACCGCTTCGTCGCGGTGGGAGTGGGACTGGCGAG	6166	
	DGKAHPTESVKPVLDTNSIWCVRV"					
CDS	5106. .5477					
	/note="ORF3"					
	/codon_start=1					
	/product="immunogenic protein"					
	/protein_id="AAC97187.1"					
	/db_xref="GI:4033732"					
	/translation="MNMSFAPTGSRSCALGLFCCSSCFCLCCPRHRPVSLAAVV					
	GGAAAPVAVVSGVTGLILSPSQSPIQIPTPSPMPPLRGLDLVFNANPDHSAPLGV					
	TRPSAPLPHVVDLPQLGPRR"					
CDS	5147. .7129					
	/note="ORF2"					
	/codon_start=1					
	/product="capsid protein"					
	/protein_id="AAC97188.1"					
	/db_xref="GI:4033733"					
	/translation="MRPRLILLFLMFLPMLPAPPQPSGRRRRSGSGSGGFWGD					
	RVDSPQFATPIHTFPFADVTAAGAGRPQRPALPGSAWRDQAPPAATSRRP					
	RTAGAAPLTAFAHPTPPVPDVGSRGAILCROYNLSTPLTSSVATGNLVLYAAPL					
	SPLLPLQDNTNTHIMATASNAQYVARAAIIRYPLVNAVGVYLSISFWPQTITT					
	PTSDMNSITSTDRVILVQPIASBELVIPSERLHYRNQGRSVETSGVAREEATSELV					
	MUCIHGSPVNSYNTPIYTGALGLDFALELEFRLNTPGNTNTRVRSYSTARRHLRRG					
	ADGTAELTAAATFRMKDLYTSTNGVGEIRGIALTNLADLILGLUPTLISAG					
	QULFYRPPVVSANGPEPTVKLYTSVENAQDQKIAI PHDIDLGESVVIQYDNQHQD					
	RPTSPAPSRPFVSLRANDVLISQAAEYDQSTGYSSTGPVYVSDTPIVNVNTAQ					
	AVRSLDWTVTLDGRPLSTIQOYSKTFVFLPURLKLSFWEAGTKAGVPYNNVTAS					
	GOLLVNAAGRVAISYVTTSLGAPVSI SAVAVLAPHSA LLEDLTLDYPABAHTFD					
	DFCPBCRPLGLQCAQFSTVAEILQRLKMKVGTREL"					
ORIGIN						

1021 CTCACACACGCTGTACCGCTTTATGAGGACCTCTATTCTACTAGTACTAATGGT 1080
Db |||||
6167 CTCACACACGCTGTACCGCTTTATGAGGACCTCTATTCTACTAGTACTAATGGT 6226
Qy |||||
Db |||||
1081 GTCGGTGAGATCGCGCGGGGATAGCCCTCACCCCTGTTCAACCTTGTCTGACACTCTGCTT 1140
Qy |||||
Db |||||
6227 GTTGGTGAGATCGCGCGGGGATAGCCCTCACCCCTGTTCAACCTTGTCTGACACTCTGCTT 6286
Qy |||||
Db |||||
1141 GCGCGCTTCCGACAGAAATGATTTGTCGCGCTGTGGCAGCTGTTCTACTCCCGTCCC 1200
Qy |||||
Db |||||
6287 GCGCGCTTCCGACAGAAATGATTTGTCGCGCTGTGGCAGCTGTTCTACTCCCGTCCC 6346
Qy |||||
Db |||||
1201 GTTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
Qy |||||
Db |||||
6347 GTTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 6406
Qy |||||
Db |||||
1261 CAGGATAAGGATTTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGTTATT 1320
Qy |||||
Db |||||
6407 CAGGATAAGGATTTGCAATCCCGCATGATTTGACCTCGGAGAAATCTCGTGTGTTATT 6466
Qy |||||
Db |||||
1321 CAGGATAAGGATTTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGTTATT 1380
Qy |||||
Db |||||
6467 CAGGATAAGGATTTGCAATCCCGCATGATTTGACCTCGGAGAAATCTCGTGTGTTATT 6526
Qy |||||
Db |||||
1381 CTTTCTCTGCTTCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCCTGCGGAGTAT 1440
Qy |||||
Db |||||
6527 CTTTCTCTGCTTCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCCTGCGGAGTAT 6586
Qy |||||
Db |||||
1441 GACCAAGTCCACTTATGGCTTCTCGACTGCGCCAGTTTATGTTTCTGACTCTCTGACCTTGT 1500
Qy |||||
Db |||||
6587 GACCAAGTCCACTTATGGCTTCTCGACTGCGCCAGTTTATGTTTCTGACTCTCTGACCTTGT 6646
Qy |||||
Db |||||
1501 GTTAAATGTCGACCGCGCGGAGCGCTGTCGCGCTGCTCGATTTGACCAAGTCTACA 1560
Qy |||||
Db |||||
6647 GTTAAATGTCGACCGCGCGGAGCGCTGTCGCGCTGCTCGATTTGACCAAGTCTACA 6706
Qy |||||
Db |||||
1561 CTTGACGCTGCGCCCTCTCCACCTCAGCAGTACTCGAAGACCTTCTTTGCTCGCGG 1620
Qy |||||
Db |||||
6707 CTTGATGCGCGCCCTCTCTCCACCTCAGCAGTACTCGAAGACCTTCTTTGCTCGCGG 6766
Qy |||||
Db |||||
1621 CTTCCGCGTAAAGTCTCTTTCTGGGAGGAGGAGCAACAACTAAAGCCGGGTACCTTTAAT 1680
Qy |||||
Db |||||
6767 CTTCCGCGTAAAGTCTCTTTCTGGGAGGAGGAGCAACAACTAAAGCCGGGTACCTTTAAT 6826
Qy |||||
Db |||||
1681 TATAACACACTGTAGAGACCAACTGCTTTGAGAAATGCGCGGACCGGTCGCT 1740
Qy |||||
Db |||||
6827 TACAAACACACTGTAGAGGAGCAACTGCTTTGAGAAATGCGCGGACCGGTTGCT 6886
Qy |||||
Db |||||
1741 ATTTCACCTTACACACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Qy |||||
Db |||||
6887 ATCTCCACTTACACACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6946
Qy |||||
Db |||||
1801 TTAGCCCCCTCTGCGCTAGCATTGCTTGGAGTACCTTGGAGTACCTTGGAGTACCTTGGAGT 1860
Qy |||||
Db |||||
6947 TTGGCCCCCTCTGCGCTAGCATTGCTTGGAGTACCTTGGAGTACCTTGGAGTACCTTGGAGT 7006
Qy |||||
Db |||||
1861 CATACTTTTATGATTTCTGCGCAGAGTGGCGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1920
Qy |||||
Db |||||
7007 CATACTTTTATGATTTCTGCGCAGAGTGGCGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 7066
Qy |||||
Db |||||
1921 CAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Qy |||||
Db |||||
7067 CAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7126
Qy |||||
Db |||||
1981 TAG 1983
Qy |||||
Db |||||
7127 TAG 7129

RESULT 15
AF185822 7143 bp RNA linear VRL 14-JUN-2001
LOCUS
DEFINITION Hepatitis E virus nonstructural polyprotein gene, partial cds;

structural protein gene, complete cds; and unknown gene.
AF185822
AF185822.1 GI:10336832
KEYWORDS
SOURCE
ORGANISM
Hepatitis E virus
Hepatitis E virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Hepatitis
E-like viruses.
REFERENCE
1 (bases 1 to 7143)
van Cuyck-Gandre,H., Zhang,H.Y., Tsarev,S.A., Warren,R.L.,
Caudill,J.D., Snellings,N.J., Begot,L., Innis,B.L. and Longer,C.F.
Short report: phylogenetically distinct hepatitis E viruses in
Pakistan
Am. J. Trop. Med. Hyg. 62 (2), 187-189 (2000)
20271579
10813471
REFERENCE
2 (bases 1 to 7143)
van Cuyck-Gandre,H., Zhang,H.Y., Tsarev,S.A., Warren,R.L.,
Caudill,J.D., Snellings,N.J., Begot,L., Innis,B.L., Longer,C.F. and
Ticehurst,J.
Direct Submission
Submitted (13-SEP-1999) Agents Transmissibles, CRSSA, 24 Ave des
Maquis du Gresivausan, BP 87, La Tronche 38702, France
Location/Qualifiers
1. 7143
/organism="Hepatitis E virus"
/mol_type="genomic RNA"
/strain="Abb-2B"
/db_xref="taxon:12461"
/country="Pakistan"
<1. 5058
/note="ORF1"
/codon_start=1
/product="nonstructural polyprotein"
/protein_id="AAG16764.1"
/db_xref="GI:10336833"
/translation="APGTTTAEQAALAAANALANAVVVRPFLSHQQTTEILINMQP
RQLVPRFVWNPFIQVTHNELEYCRASGRGLEIGHPRCSINPNVHRCFLRP
VGRDVQWYATPTGPAACRRSALRGLPAADRTYCFDGFSGCSFPAAGIALYSLHD
MPSDVAAMFHMTRLYAALHLPPEVLLPGTTRTASLYLLIHGRRVVVYEGDTS
AGYNDVNLRSWIKTKVTGDHPLVIERVRAIGHFVLLLTAAPEPSPMPVYPRPS
TEVYRSIFGCGTSLPFTSCSTSTFAVPAHIDRLMLFGATLDDQAFCCSRILMT
YLRGISYKVTGTLVANEGWASDALTAIVTAAYLTICHQRYLRTQALVKQWRILIER
EHAOKFIRLXSMLEPGRDVPCGRLEFYAQRRLSAGPHLDPRVLVPDESALSD
CETAIRKALSFCCEKMWLGOCCTCLOPAEGVGDQGDNEAYSGSDVDPAESALSD
ISGSDVPEALQPLDQALDPAEIVARAGRLTATVKVSDVGRIDCEITLLGNKFTI
SFVDGAVLETNGFERHNSFDASQSTMAAGPFLTYAASAGLEVRYVVAAGLDHRAVF
APGVSAPAGEVTAFCSALYFNREARHSLTGNLWFEGLLGLFAPFSPGHVWES
ANPFCGESLTTRTWSEVDVSSPARPDIGFMSEPSISRAATPTLAAPLSPAPDLS
TPSPAPDEPASGATAGAPATTHQAAHRRLLTTPDGSKVPAGSLPSTCTWLVA
SNADHRPGGLCHAFQVYPASFDAAAFVMDGAAATLTTPRPIHVAVDYLRLEHNP
KRLEAARETCSRLGTAAYPLLTGTIYQVPGSPDAMERNHRPDELAYLPELAARWF
ENRTPRALTITTEDVARTANLALBELDSATDVRACGRVTPGVVQVQFTAGVPGSG
KRSRTQADVDVVVTRELNARMRRGFAFPTHTAARVTQGRVWIDEAPSLPPLH
LILLHRAATVHLGLDNPQIPADIDFAGLVPALPDPLATSMWVTHRCADVCELI
RGAYPMIOTTSRLSLWGEPAVGOKLVFOAKAANPGSVTHHEAOCATYTTETII
ATADAGLIQSRRAHAIWALTPHTEKCVIIDAPGLLREVGISDAIVNPFLLAGLEIGH
QRPVSI PRGNPDANDVTILAAPPPCQISAFHQLABELGHRPAPVAALPPCPLEBGL
LYLPELITTCDSVTFELTDIVHCRMAAPSRQKAVLSTLVGRITKLYFNASHSDVR
DSLARFI PAIGEVQVTCELYELVEAMVEKQDGSVAVLELDLCNRDVSRTIPFKDCN
KFTTGETTAGHVGQGI SAWSKTF CALGPWFRAIEKATLALLPGCVFGDADTFV
SAAVAAKASAMVFNDFEOSTNNFSLGLECAIMECGMPQWILIRLHLSRAWIL
CAPKESLRGPKKHSGEFTLWNTVMNAVTTCYDFRDLQVAAFGKGDLSIVLSEY
RQSPGAUVLIACGLKLVDFRPIGIGYAGVVVAPGLGALPVDVVRAGRLTEKNWGPFP
ERAQRLAVSDFLRKLTNVAQMCVDVVSVRVYGVSPGLVHNLIGMLQAVADGKAHFTF
SVKPELDLTNSILCRVE"
5055. 5426
/note="ORF3"
/codon_start=1
/product="unknown"
/protein_id="AAG16765.1"
/db_xref="GI:10336834"
/translation="MNNWSFAAPMGSRPCALGLFCCSSCFCLCCPRHRPVSRLAADV

Qy	1801	TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGGA	CTTACCTGCCCGCGCC	1860
Db	6896	TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGGA	CTTACCTGCCCGCGCC	6955
Qy	1861	CATACTTTTGATGATTTCTGCCCCAGAGTGCGCCCCCTTGCC	CTTCAGGGCTGCGCTTTC	1920
Db	6956	CATACTTTTGATGATTTCTGCCCCAGAGTGCGCCCCCTTGCC	CTTCAGGGCTGCGCTTTC	7015
Qy	1921	CAGTCTACTGCTGAGCTTCAGCGCCTTAAGATGAAGTGGTAA	AACTCGGGAGTTG	1980
Db	7016	CAGTCTACTGCTGAGCTTCAGCGCCTTAAGATGAAGTGGTAA	AACTCGGGAGTTA	7075
Qy	1981	TAG	1983	
Db	7076	TAG	7078	

Search completed: August 21, 2004, 15:14:59
Job time : 7653.28 secs

Blank sheet

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 08:40:47 ; Search time 765.579 Seconds
(without alignments)
11003.666 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129
Perfect score: 1983
Sequence: 1 ATGCGCCCTCGGCTATTATTT.....GTAAACTCGGAGTGTAG 1983

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1983	100.0	1983	2	AAT96959 Hepatitis
2	1983	100.0	2049	2	AAT27107 Hepatitis
3	1983	100.0	2094	2	AAT27107 Hepatitis
4	1983	100.0	2094	2	AAT27107 Hepatitis
5	1983	100.0	2094	2	AAT27107 Hepatitis
6	1983	100.0	2094	2	AAT27107 Hepatitis
7	1983	100.0	2094	2	AAT27107 Hepatitis
8	1983	100.0	2094	2	AAT27107 Hepatitis
9	1983	100.0	2094	2	AAT27107 Hepatitis
10	1983	100.0	2094	2	AAT27107 Hepatitis
11	1970.2	99.8	7195	9	AD224374
12	1949.4	98.3	7195	2	AAV61688
13	1949.4	98.3	7194	2	AAV61688
14	1805.4	91.0	1984	2	AAT27107 Hepatitis
15	1805.4	91.0	1990	6	ABL60053
16	1805.4	91.0	1990	9	AD224374
17	1805.4	91.0	2054	4	AAV61688
18	1793.4	90.4	1989	6	ABL60054
19	1793.4	90.4	1989	9	AD224374
20	1791	90.3	7204	8	ADA50063
21	1789.4	90.2	7204	8	ADA50063
22	1789.4	90.2	7204	8	ADA50063
23	1787.8	90.2	7158	2	AAT27394 Hepatitis

24	1787.8	90.2	7168	2	AAQ45197
25	1787.8	90.2	7168	2	AAV71604
26	1787.8	90.2	7204	8	ADA50064
27	1647	83.1	1647	2	AAT27109
28	1617.2	81.6	1647	2	AAV61686
29	1506.4	76.0	1643	2	AAQ27436
30	1384.6	69.8	2058	2	AAT27108
31	1368.6	69.0	1980	2	AAT96960
32	1368.6	69.0	2100	2	AAQ47130
33	1368.6	69.0	7171	2	AAQ14413
34	1368.6	69.0	7171	2	AAV66322
35	1368.6	69.0	7171	2	AAV54730
36	1368.6	69.0	7171	3	AAV54730 Composite
37	1368.6	69.0	7171	4	AAV54730 Composite
38	1368.6	69.0	7171	4	AAV54730 Composite
39	1368.6	69.0	7171	6	AAV54730 Composite
40	1368.6	69.0	7180	9	AD224378
41	1367.6	69.0	7230	9	AD224378
42	1367.6	69.0	7230	9	AD224378
43	1347.6	68.0	7228	9	AD224378
44	1346.8	67.9	7233	9	AD224378
45	1346	67.9	7234	9	AD224378

ALIGNMENTS

RESULT 1
AAT96959
ID AAT96959 standard; DNA; 1983 BP.
XX
AC AAT96959;
XX
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 26-FEB-1998 (first entry)
XX
DE Hepatitis E virus Burma strain ORF2.
XX
KW Hepatitis E virus; Burma; Mexico; immunoassay; peptide antigen; antibody;
KW diagnosis; HEV; ds.
XX
OS Hepatitis E virus; - Burma strain.
XX
FH Key Location/Qualifiers
FT CDS l..1983
FT /*tag= a
FT /product= "ORF2_protein"
XX
PN US5686239-A.
XX
PD 11-NOV-1997.
XX
PF 09-MAY-1994; 94US-00240049.
XX
PR 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-0036672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 17-JAN-1991; 91US-00681078.
PR 20-APR-1992; 92US-00822335.
PR 01-MAY-1992; 92US-00870985.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Tam AW, Reyes GR, Yarbough PO;
XX
DR WPI; 1997-558132/51.
XX
P-PSDB; AAW35826.
PT Diagnosis of hepatitis E virus Burma and Mexico strain infection - by

PT immunoassay with hepatitis E virus peptide antigens.

XX PS Disclosure; Col 23-24; 36pp; English.

XX CC A method has been developed for detecting hepatitis E virus (HEV)
XX CC antibodies (Ab). The method comprises: (a) reacting a serum sample with a
XX CC HEV peptide antigen; and (b) examining the peptide for the presence of
XX CC bound Ab, where the presence of bound Ab indicates the presence of HEV
XX CC Ab. The present sequence encodes the protein from the open reading frame,
XX CC ORF2, from HEV Burma strain. The method can be used to diagnose infection
XX CC with the enterically transmitted non-A/non-B viral hepatitis agent HEV,
XX CC specifically the HEV Burma and Mexico strains. (Updated on 25-MAR-2003 to
XX CC correct PF field.) (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 1983 BP; 340 A; 665 C; 486 G; 492 T; 0 U; 0 Other;

Query Match 100.0%; Score 1983; DB 2; Length 1983;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTGGCTATGCTGCGCGCCA	60
DB	1	ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTGGCTATGCTGCGCGCCA	60
QY	61	CGCGCGGTCAGCGCTCGCGCGCTGCTGCGCGCGCGAGCGCGGTTCCGCGGTGT	120
DB	61	CGCGCGGTCAGCGCTGCGCGCGCTGCTGCGCGCGCGAGCGCGGTTCCGCGGTGT	120
QY	121	TTCTGCGGTCAGCGGTCGATTTCTAGCCCTTCGCAATCCCTCATATTCATCAACCAAC	180
DB	121	TTCTGCGGTCAGCGGTCGATTTCTAGCCCTTCGCAATCCCTCATATTCATCAACCAAC	180
QY	181	CCCTGCGCCCGGATGTCAGCGCTGCGCGCGGCTGAGCTGTTGCGCAACCGGC	240
DB	181	CCCTGCGCCCGGATGTCAGCGCTGCGCGCGGCTGAGCTGTTGCGCAACCGGC	240
QY	241	CGACCACTCGGCTCGCTGCGGTCGACAGCGCCAGCGCGCGCTGCTCAACGTCGT	300
DB	241	CGACCACTCGGCTCGCTGCGGTCGACAGCGCCAGCGCGCGCTGCTCAACGTCGT	300
QY	301	AGACCTACCAAGTCGCGGCGCGCTAACCGCGGTGCTCGCGCGCATGACACCCG	360
DB	301	AGACCTACCAAGTCGCGGCGCGCTAACCGCGGTGCTCGCGCGCATGACACCCG	360
QY	361	CGAGTGCATGTCGACCTCCGCGGCGCATCTTGGCGCGGAGTATACCTATCAACA	420
DB	361	CGAGTGCATGTCGACCTCCGCGGCGCATCTTGGCGCGGAGTATACCTATCAACA	420
QY	421	TCCTCCCTTACCTTCCGCGGCGCACCGGCACTAACTGCTTCTTATGCGCGCTCT	480
DB	421	TCCTCCCTTACCTTCCGCGGCGCACCGGCACTAACTGCTTCTTATGCGCGCTCT	480
QY	481	AGTCCGCTTTACCCCTTCAGGACGGCACCAATACCATATAATGCGCACGGAAGCTTCT	540
DB	481	AGTCCGCTTTACCCCTTCAGGACGGCACCAATACCATATAATGCGCACGGAAGCTTCT	540
QY	541	AATATGCGGCTACCGGTTGCGCGGTCGACATCCGTTACCGCGCGCTGCTCCCAAT	600
DB	541	AATATGCGGCTACCGGTTGCGCGGTCGACATCCGTTACCGCGCGCTGCTCCCAAT	600
QY	601	GCTGTGCGGCTTACGCACTCTCCATCTCTTCTGCGCACAGACCAACACCCGACG	660
DB	601	GCTGTGCGGCTTACGCACTCTCCATCTCTTCTGCGCACAGACCAACACCCGACG	660
QY	661	TCCGTTGATATGAATTAATAAAGTTCGATATTTAGTTCAGCGCGGCATA	720
DB	661	TCCGTTGATATGAATTAATAAAGTTCGATATTTAGTTCAGCGCGGCATA	720
QY	721	GCCTCTGAGCTTGATCCCAAGTGAGCGCTACACTATGTAACCAAGGCTGCGCTCC	780
DB	721	GCCTCTGAGCTTGATCCCAAGTGAGCGCTACACTATGTAACCAAGGCTGCGCTCC	780
QY	781	GTCGAGACCTCTGCGGTGGCTGAGGAGGAGCTACCTCTGCTTGTATGCTTTGCATA	840

DB	781	GTCGAGACCTCTGCGGTGGCTGAGGAGGAGCTACCTCTGCTTGTATGCTTTGCATA	840
QY	841	CATGCTCACTCGTAAATTCCTATATAACACCTATACCGTGCCTCGGCTGTTG	900
DB	841	CATGCTCACTCGTAAATTCCTATATAACACCTATACCGTGCCTCGGCTGTTG	900
QY	901	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACGCGGTC	960
DB	901	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACGCGGTC	960
QY	961	TCCGCTTATTCAGACACTCTCGCAACCGCTTGTGCGGTGCGGACGCGGCTGCGGAG	1020
DB	961	TCCGCTTATTCAGACACTCTCGCAACCGCTTGTGCGGTGCGGACGCGGCTGCGGAG	1020
QY	1021	CTCACCAACACCGCTGCTACCGCTTATGAGGAGCTCTATTTACTAGTACTAATGTT	1080
DB	1021	CTCACCAACACCGCTGCTACCGCTTATGAGGAGCTCTATTTACTAGTACTAATGTT	1080
QY	1081	GTCGCTGAGATCGGCGCGGATAGCCCTCACCCCTGTTCAACCTTCTGACACTGCTT	1140
DB	1081	GTCGCTGAGATCGGCGCGGATAGCCCTCACCCCTGTTCAACCTTCTGACACTGCTT	1140
QY	1141	GCGCGCTCGCACAGAAATTGCTGCGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTT	1200
DB	1141	GCGCGCTCGCACAGAAATTGCTGCGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTT	1200
QY	1201	GTTGCTCAGCAATAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1260
DB	1201	GTTGCTCAGCAATAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1260
QY	1261	CAGGATTAAGGCTTATGCAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1320
DB	1261	CAGGATTAAGGCTTATGCAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1320
QY	1321	CAGGATTAAGGCTTATGCAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1380
DB	1321	CAGGATTAAGGCTTATGCAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1380
QY	1381	CCTTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1440
DB	1381	CCTTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1440
QY	1441	GACAGTCCACTTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1500
DB	1441	GACAGTCCACTTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1500
QY	1501	GTTAATGTTGCGACCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1560
DB	1501	GTTAATGTTGCGACCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1560
QY	1561	CTTGAAGTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1620
DB	1561	CTTGAAGTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1620
QY	1621	CTCGCGGTAAGCTCTCTTCTGCGGAGGACAGCAAACTAAAGCCGGTACCTTATAAT	1680
DB	1621	CTCGCGGTAAGCTCTCTTCTGCGGAGGACAGCAAACTAAAGCCGGTACCTTATAAT	1680
QY	1681	TATAACCACTGCTAGCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1740
DB	1681	TATAACCACTGCTAGCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	1740
QY	1741	ATTTCACCTTACCACTAGCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTT	1800
DB	1741	ATTTCACCTTACCACTAGCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTT	1800
QY	1801	TTAGCCCCCACTCTGCGCTAGCATCTTGAAGATACCTTGAGTACCTGCGCGGCC	1860
DB	1801	TTAGCCCCCACTCTGCGCTAGCATCTTGAAGATACCTTGAGTACCTGCGCGGCC	1860
QY	1861	CATACCTTCTGATGATTTCTGCGGAGTGGCGGCTGCTGCTGCTGCTGCTGCTT	1920


```

Sequence 2094 BP; 353 A; 697 C; 508 G; 536 T; 0 U; 0 Other;
Very Match 100.0%; Score 1983; DB 2; Length 2094;
Est Local Similarity 100.0%; Pred. No. 0;
tches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	61	CCGCCCGGGTCAGCCGCTTGACCCCGGTCGGTGGCGGCGGTCGCGCGGGTGGT	120
Db	106	CCGCCCGGGTCAGCCGCTTGACCCCGGTCGGTGGCGGCGGTCGCGCGGGTGGT	165
QY	121	TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTATCCCAACCAAC	180
Db	166	TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTATCCCAACCAAC	225
QY	181	CCCTTCGCCCCCGATGTACCGTGCGGCGGGCTGGACCTCGTGTTCGCCCAACCCGCC	240
Db	226	CCCTTCGCCCCCGATGTACCGTGCGGCGGGCTGGACCTCGTGTTCGCCCAACCCGCC	285
QY	241	GGACCACTCGGCTCCGCTTGGCTGACAGGCCACGCGCCCGCTTGGCTCAGCTCGT	300

XX	HEV third and second ORFs.
DE	
XX	Enterically transmitted non-A non-B hepatitis; ET-NANB; vaccine; ss.
KW	
XX	
OS	Hepatitis E virus; Burma strain.
OS	
XX	Key
FH	Location/Qualifiers
XX	

Db 286 CGACCACTCGGCTCGGCTTGGGTGACCAAGGCCAGCGCCCGCGTTCGCTCAAGTCGT 345
Qy 301 AGACCTACCAAGCTGGGCGCGCGCTAAACGGGGTGGCTCCGCGCCCATGACACCCCG 360
Db 346 AGACCTACCAAGCTGGGCGCGCGCTAAACGGGGTGGCTCCGCGCCCATGACACCCCG 405
Qy 361 CCAGTGCCTGATGTCGACTCCGCGGCGGCATCTTTGGCGCGGAGTATAACCTATCAACA 420
Db 406 CCAGTGCCTGATGTCGACTCCGCGGCGGCATCTTTGGCGCGGAGTATAACCTATCAACA 465
Qy 421 TCTCCGCTTACCTCTTCGTGGCGACCGGCACCTAACCTGGTTCCTTTATGCGCGCCCTCTT 480
Db 466 TCTCCGCTTACCTCTTCGTGGCGACCGGCACCTAACCTGGTTCCTTTATGCGCGCCCTCTT 525
Qy 481 AGTCCGCTTTTACCCCTTCAGGACCGCACCAATACCATATAATGSCCAGCAAGCTTCT 540
Db 526 AGTCCGCTTTTACCCCTTCAGGACCGCACCAATACCATATAATGSCCAGCAAGCTTCT 585
Qy 541 AATTATGCCAGTACCGGGTTGCCGTGCCCAATCCGTTACCGCCCGCTGGTCCCCAT 600
Db 586 AATTATGCCAGTACCGGGTTGCCGTGCCCAATCCGTTACCGCCCGCTGGTCCCCAT 645
Qy 601 GCTGTGGGGTTAGCCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCGAGC 660
Db 646 GCTGTGGGGTTAGCCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCGAGC 705
Qy 661 TCCGTTGATATGAATTCATTAACCTCGACGGATGTTGTTATTTAGTCCAGCCCGCAT 720
Db 706 TCCGTTGATATGAATTCATTAACCTCGACGGATGTTGTTATTTAGTCCAGCCCGCAT 765
Qy 721 GCCTCTGAGCTTGTGATCCCAAGTGAAGCCCTACACTATCGTAACCAAGCTGGCGCTCC 780
Db 766 GCCTCTGAGCTTGTGATCCCAAGTGAAGCCCTACACTATCGTAACCAAGCTGGCGCTCC 825
Qy 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACTCTGGTCTTGTATGCTTTGCATA 840
Db 826 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACTCTGGTCTTGTATGCTTTGCATA 885
Qy 841 CATGCTCACTCGTAAATTCCTATATAACACCTATACCGGTAAACCAATACGGGGTGG 900
Db 886 CATGCTCACTCGTAAATTCCTATATAACACCTATACCGGTAAACCAATACGGGGTGG 945
Qy 901 GACTTTGGCCCTTGAGCTTGAGTTTGGCAACCTTACCCCGGTAAACCAATACGGGGTGG 960
Db 946 GACTTTGGCCCTTGAGCTTGAGTTTGGCAACCTTACCCCGGTAAACCAATACGGGGTGG 1005
Qy 961 TCCCGTTATTCAGCACTGCTGCCACCGCTTCGTGGCGTGGGACGGGACTGCCGAG 1020
Db 1006 TCCCGTTATTCAGCACTGCTGCCACCGCTTCGTGGCGTGGGACGGGACTGCCGAG 1065
Qy 1021 CTCACCAACCGGCTGTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGCT 1080
Db 1066 CTCACCAACCGGCTGTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGCT 1125
Qy 1081 GTCGCTGAGATGGCGCGGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTT 1140
Db 1126 GTCGCTGAGATGGCGCGGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTT 1185
Qy 1141 GCGCGCTTCCGACAGAAATGATTTGCTGGGTGGTGGCAGCTGTTCTACCTCCGTCGC 1200
Db 1186 GCGCGCTTCCGACAGAAATGATTTGCTGGGTGGTGGCAGCTGTTCTACCTCCGTCGC 1245
Qy 1201 GTTGTCTAGCAATGGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAACTCTCAG 1260
Db 1246 GTTGTCTAGCAATGGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAACTCTCAG 1305
Qy 1261 CAGGATAAGGTTATGCAATCCCGCATGACATTTGACCTCGGAGAACTCTCGTGGTTATT 1320
Db 1306 CAGGATAAGGTTATGCAATCCCGCATGACATTTGACCTCGGAGAACTCTCGTGGTTATT 1365
Qy 1321 CAGGATTATGATAACCAATGAAACAGATCGGCGGAGCGCTTCTCCAGCCCATCGCGC 1380
Db 1366 CAGGATTATGATAACCAATGAAACAGATCGGCGGAGCGCTTCTCCAGCCCATCGCGC 1425

Qy 1381 CCTTTCTGTCTTCGAGCTAATGATGCTTTTGGCTCTCTCTCACCGCTGCCGAGTAT 1440
Db 1426 CCTTTCTGTCTTCGAGCTAATGATGCTTTTGGCTCTCTCTCACCGCTGCCGAGTAT 1485
Qy 1441 GACCACTCCACTTATGGCTCTTCGACTGGCCAGATTTATGTTTCTGACTCTGTGACCTTG 1500
Db 1486 GACCACTCCACTTATGGCTCTTCGACTGGCCAGATTTATGTTTCTGACTCTGTGACCTTG 1545
Qy 1501 GTTAATGTTGACCGGCGCGAGCGGTTGCCCGTCTCGATTTGGAACCAAGGTCA 1560
Db 1546 GTTAATGTTGACCGGCGCGAGCGGTTGCCCGTCTCGATTTGGAACCAAGGTCA 1605
Qy 1561 CTTGACGGTGGCCCTCTCCACCACTCCAGCAGTACTCGAAGACCTTTCTTTGCTCCGCG 1620
Db 1606 CTTGACGGTGGCCCTCTCCACCACTCCAGCAGTACTCGAAGACCTTTCTTTGCTCCGCG 1665
Qy 1621 CTTCCGGTAACTCTCTTTCTGGAGGAGGACCAACTAAAGCCGGGTACCTTTATAT 1680
Db 1666 CTTCCGGTAACTCTCTTTCTGGAGGAGGACCAACTAAAGCCGGGTACCTTTATAT 1725
Qy 1681 TATAACCACTGCTAGGACCAACTGCTTTGCGAGAAATGCCCGCGGACCCGGTCTGCT 1740
Db 1726 TATAACCACTGCTAGGACCAACTGCTTTGCGAGAAATGCCCGCGGACCCGGTCTGCT 1785
Qy 1741 ATTTCACCTTACCACTAGCTGCTGGTCTGGTCCGCTCTCCATTTCTGCGGTTGCCGTT 1800
Db 1786 ATTTCACCTTACCACTAGCTGCTGGTCTGGTCCGCTCTCCATTTCTGCGGTTGCCGTT 1845
Qy 1801 TTAGCCCCCACTCTGCGCTAGCACTTGTGAGGATACCTTGGACTACCTGCCCGCGCC 1860
Db 1846 TTAGCCCCCACTCTGCGCTAGCACTTGTGAGGATACCTTGGACTACCTGCCCGCGCC 1905
Qy 1861 CATACTTTGATGATTTCTGCCAGAGTCCGCCCTTTGGCTTCAGGGCTGGCGTTTC 1920
Db 1906 CATACTTTGATGATTTCTGCCAGAGTCCGCCCTTTGGCTTCAGGGCTGGCGTTTC 1965
Qy 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAGATGAAGGTGGTAAACTCGGAGTTG 1980
Db 1966 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAGATGAAGGTGGTAAACTCGGAGTTG 2025
Qy 1981 TAG 1983
Db 2026 TAG 2028
RESULT 4
AAQ46813
ID AAQ46813 standard; DNA; 2094 BP.
XX AC AAQ46813;
XX 25-MAR-2003 (revised)
DT 14-FEB-1994 (first entry)
XX XX
XX Burma strain HEV ORF2 and ORF3.
KW Hepatitis E virus; vaccine; neutralising antibodies; infection; block;
open reading frame; antibodies; ss.
XX OS Hepatitis E virus.
XX Key Location/Qualifiers
FT CDS 5..311
FT /tag= a
FT /note= "ORF2"
FT 46..2094
FT /tag= b
FT /note= "ORF3"
XX WO9314208-A2.
XX 22-JUL-1993.
PD

XX 19-JAN-1993; 93WO-US000475.
 XX PF 17-JAN-1992; 92US-00822335.
 XX PR 20-APR-1992; 92US-00870985.
 XX PA (GENE-) GENELABS TECHNOLOGIES INC.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Reyes GR, Bradley DW, Twu J, Purdy MA, Tam AM, Krawczynski KZ;
 XX DR WPI; 1993-243223/30.
 XX DR P-PSDB; AAR39305, AAR39306.
 XX PT Antigen and antibody vaccines against hepatitis B virus infection -
 PT contain peptide(s) derived from capsid protein C-terminal or antibodies
 PT against protein.
 XX PS Disclosure; Fig 7; 43pp; English.
 XX CC The sequence is that encoding Burma strain hepatitis B virus (HEV) open
 CC reading frames ORF2 and ORF3. ORF2 encodes the putative capsid protein of
 CC the virus, this protein or peptide fragments of it may be used in a
 CC vaccine composition for immunising an individual against HEV. Antibodies
 CC raised against these peptides can also be used in such vaccines. (Updated
 CC on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 2094 BP; 353 A; 697 C; 508 G; 536 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1983; DB 2; Length 2094;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCGCGGCCA 60
 DB 46 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCGCGGCCA 105
 QY 61 CCGCCCGGTGACGGCTGCGCGCGTCTGCGCGGCGCGCGCGGCTTCCGCGGTGGT 120
 DB 106 CCGCCCGGTGACGGCTGCGCGCGTCTGCGCGGCGCGCGCGGCTTCCGCGGTGGT 165
 QY 121 TTCTGGGTGACCGGGTGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
 DB 166 TTCTGGGTGACCGGGTGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 225
 QY 181 CCCTTCCGCCCGGATGTACGGCTGCGCGCGGCTGCGCGGCTGCGCGGCTGCGCGGCT 240
 DB 226 CCCTTCCGCCCGGATGTACGGCTGCGCGCGGCTGCGCGGCTGCGCGGCTGCGCGGCT 285
 QY 241 CGACCACTCGGCTCCGCTTGGCTGACCGCGGCTGACCGCGGCTGACCGCGGCTGAC 300
 DB 286 CGACCACTCGGCTCCGCTTGGCTGACCGCGGCTGACCGCGGCTGACCGCGGCTGAC 345
 QY 301 AGACCTACCACTGCGGCGCGGCTGCGCGGCTGCGCGGCTGCGCGGCTGCGCGGCT 360
 DB 346 AGACCTACCACTGCGGCGCGGCTGCGCGGCTGCGCGGCTGCGCGGCTGCGCGGCT 405
 QY 361 CGAGTGCCTGATGTGCGACTCCGCGCGGCTGCGCGGCTGCGCGGCTGCGCGGCTGCA 420
 DB 406 CGAGTGCCTGATGTGCGACTCCGCGCGGCTGCGCGGCTGCGCGGCTGCGCGGCTTCA 465
 QY 421 TCTCCCTTACCTCTTCCGTTGGCCACCGGCACTAACTGCTTCTTTATGCGCGCCCTCT 480
 DB 466 TCTCCCTTACCTCTTCCGTTGGCCACCGGCACTAACTGCTTCTTTATGCGCGCCCTCT 525
 QY 481 AGTCCGCTTTTACCCCTTCAGGACCGGCAACCAATACCATATATATGCGGCAAGCTTCT 540
 DB 526 AGTCCGCTTTTACCCCTTCAGGACCGGCAACCAATACCATATATATGCGGCAAGCTTCT 585
 QY 541 AATTATGCCAGTACCGGGTTCGCGGTCGCGGTCGCGGTCGCGGTCGCGGTCGCGGTC 600
 DB 586 AATTATGCCAGTACCGGGTTCGCGGTCGCGGTCGCGGTCGCGGTCGCGGTCGCGGTC 645

QY 601 GCTGTGCGGGTTAGCCCATCTCCATCTCATTCTGCGCCACAGACACACACACCCCGACG 660
 DB 646 GCTGTGCGGGTTAGCCCATCTCCATCTCATTCTGCGCCACAGACACACACCCCGACG 705
 QY 661 TCCGTTGATATGAATTCATAAATCAATCACTCGACGGATGTTCTGATATTTTATGTCAGCCGCGGATA 720
 DB 706 TCCGTTGATATGAATTCATAAATCAATCACTCGACGGATGTTCTGATATTTTATGTCAGCCGCGGATA 765
 QY 721 GCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGGCTGCGGCTCC 780
 DB 766 GCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGGCTGCGGCTCC 825
 QY 781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGGTCTTGTGTTATGCTTTGCATA 840
 DB 826 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGGTCTTGTGTTATGCTTTGCATA 885
 QY 841 CATGGCTCACTCGTAAATTCCTATATACCTATACCTATACCGGTGCGGCTGCGGCTGTTG 900
 DB 886 CATGGCTCACTCGTAAATTCCTATATACCTATACCTATACCGGTGCGGCTGCGGCTGTTG 945
 QY 901 GACTTTGGCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTC 960
 DB 946 GACTTTGGCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTC 1005
 QY 961 TCCGCTTATTCAGCACTGCTCGCCACCGCTTCTGTCGCGGTGCGGACGCGGACTGCGGAG 1020
 DB 1006 TCCGCTTATTCAGCACTGCTCGCCACCGCTTCTGTCGCGGTGCGGAGCGGACTGCGGAG 1065
 QY 1021 CTCACCAACCGCTGCTACCGCTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080
 DB 1066 CTCACCAACCGCTGCTACCGCTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1125
 QY 1081 GTCGCTGAGTACCGCGCGGATAGCCCTCACCTGTTCAACCTGCTGACACTCTGCTT 1140
 DB 1126 GTCGCTGAGTACCGCGCGGATAGCCCTCACCTGTTCAACCTGCTGACACTCTGCTT 1185
 QY 1141 GCGCGCTGCGGACAGAAATGATTCTGCTGGCTGGTGGCGAGCTGTTTCTACTCCCGTCCC 1200
 DB 1186 GCGCGCTGCGGACAGAAATGATTCTGCTGGCTGGTGGCGAGCTGTTTCTACTCCCGTCCC 1245
 QY 1201 GTTGTCTCAGCCAAATGCGGAGCGGACTGTTAAAGTTGTATACATCTCTAGAGAAATGCTCAG 1260
 DB 1246 GTTGTCTCAGCCAAATGCGGAGCGGACTGTTAAAGTTGTATACATCTCTAGAGAAATGCTCAG 1305
 QY 1261 CAGGATAAGGGTATTCGAATCCGCGATGACATTCAGCTCGGAGAAATCTCTGTGTTATT 1320
 DB 1306 CAGGATAAGGGTATTCGAATCCGCGATGACATTCAGCTCGGAGAAATCTCTGTGTTATT 1365
 QY 1321 CAGGATTATGATAACCAACATGAACAGATCGGCGGACGCTTCTCAGCCCCATCGCGC 1380
 DB 1366 CAGGATTATGATAACCAACATGAACAGATCGGCGGACGCTTCTCAGCCCCATCGCGC 1425
 QY 1381 CTTTCTCTGCTTCTCGAGCTAATGATGCTGTTTGGCTCTCTCTCAGCGTGGCGAGTAT 1440
 DB 1426 CTTTCTCTGCTTCTCGAGCTAATGATGCTGTTTGGCTCTCTCTCAGCGTGGCGAGTAT 1485
 QY 1441 GACCACTGCACTTATGCTCTTCTGAGTGGCGGAGTATGTTTCTGACTCTGTGACCTTG 1500
 DB 1486 GACCACTGCACTTATGCTCTTCTGAGTGGCGGAGTATGTTTCTGACTCTGTGACCTTG 1545
 QY 1501 GTTAAATGTTGCGACCGCGCGGAGGCTTGGCCGCTGCGTTCGATTGGACCAAGGTCA 1560
 DB 1546 GTTAAATGTTGCGACCGCGCGGAGGCTTGGCCGCTGCGTTCGATTGGACCAAGGTCA 1605
 QY 1561 CTTGAGGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCGG 1620
 DB 1606 CTTGAGGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCGG 1665
 QY 1621 CTCGCGGTGAAGCTCTCTTTCTGGGAGGAGGCAACAACTAAAGCCGGGTACCTCTATAAT 1680
 DB 1666 CTCGCGGTGAAGCTCTCTTTCTGGGAGGAGGCAACAACTAAAGCCGGGTACCTCTATAAT 1725
 QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTTCGAGAAATCCGCGCGGCGACCGGGTCTGCT 1740

```
Db 1726 TATAACACCACTGCTAGCGACCACTGCTTGTGAGAAATGCGCGGACCGGGTCGCT 1785
QY 1741 ATTTCCACTTACACCACTAGCTGCTGGTGGTCCGCTCTCCATTTCTGCGGTGCGGT 1800
Db 1786 ATTTCCACTTACACCACTAGCTGCTGGTGGTGGTCCGCTCTCCATTTCTGCGGTGCGGT 1845
QY 1801 TTAGCCCCCACTCTCGCTAGCATTTGCTTGGAGATACCTTGGACTACCTGCGCGCGC 1860
Db 1846 TTAGCCCCCACTCTCGCTAGCATTTGCTTGGAGATACCTTGGACTACCTGCGCGCGC 1905
QY 1861 CATACCTTTGATTTCTGCGGAGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1920
Db 1906 CATACCTTTGATTTCTGCGGAGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1965
QY 1921 CAGTCTACTGCTGCTAGCTTACGCGCTTAAAGATGAAGTGGTAAACTCGGGAGTTG 1980
Db 1966 CAGTCTACTGCTGCTAGCTTACGCGCTTAAAGATGAAGTGGTAAACTCGGGAGTTG 2025
QY 1981 TAG 1983
Db 2026 TAG 2028

RESULT 5
AAV66321
ID AAV66321 standard; DNA; 7195 BP.
XX
AC AAV66321;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 23-DEC-1998 (first entry)
XX
DE ET-NANB (HEV) Burma strain DNA sequence.
XX
KW Enterically transmitted non A non B hepatitis virus; ET-NANB;
KW Hepatitis E virus; HEV; Burma HEV isolate; vaccine; diagnostic probe; ds.
XX
OS Non-A.
OS non-B hepatitis virus.
XX
FH Key Location/Qualifiers
FT CDS 28..5109
FT /tag= a
FT /note= " encodes AAW80196"
FT CDS 5106..5477
FT /tag= c
FT /note= " encodes AAW80198"
FT CDS 5147..7129
FT /tag= b
FT /note= " encodes AAW80197"
XX
US5824649-A.
XX
20-OCT-1998.
XX
07-JUN-1995; 95US-00475807.
XX
17-JUN-1988; 88US-00208997.
XX
11-APR-1989; 89US-00336672.
XX
16-JUN-1989; 89US-00367486.
XX
13-OCT-1989; 89US-00420921.
XX
05-APR-1990; 90US-00505888.
XX
25-JUL-1994; 94US-00279823.
XX
(GENE-) GENELABS TECHNOLOGIES INC.
XX
Bradley DW, Reyes GR, Krawczynski KZ, Tam A, Fry KE, Yarbrough PO;
XX
WPI: 1998-582599/49.
XX
P-PSDB; AAW80196, AAW80197, AAW80198.
XX
```

```
PT Hepatitis E virus proteins - useful for diagnosis or vaccine production
XX the virus.
PS Claim 1; Col 51-58; 47pp; English.
XX
CC The present sequence represents the genome of the Burma strain of
CC enterically transmitted non A non B hepatitis virus (ET-NANB) (hepatitis
CC E virus (HEV)). The specification describes an isolated protein which is
CC specifically immunoreactive with antibodies present in individuals
CC infected with HEV and encoded by a sequence contained in an open reading
CC frame (ORF) of an HEV genome. The genome has a sequence that is more than
CC 70% identical to the ORF1 sequence from Burma HEV isolate. The protein is
CC used as a vaccine and a diagnostic probe for ET-NANB. (Updated on 25-MAR-
CC 2003 to correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 7195 BP; 1221 A; 2312 C; 1868 G; 1794 T; 0 U; 0 Other;
Query Match 100.0%; Score 1983; DB 2; Length 7195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCTATTTTGTGTGCTCTCTCAATTTTTCCTTATGCTGCGCGCCA 60
Db 5147 ATGCGCCCTCGGCTATTTTGTGTGCTCTCTCAATTTTTCCTTATGCTGCGCGCCA 5206
QY 61 CGCGCGGTCAGCGCTGCGCGCGCTGCTGGCGCGCGCGCGCGCGCGCGCGTTCGGCGTGT 120
Db 5207 CGCGCGGTCAGCGCTGCGCGCGCTGCTGGCGCGCGCGCGCGCGCGCGTTCGGCGTGT 5266
QY 121 TTCTGGGCTGACCGGGTGTGATTTCTCAGCCCTTTCGCAATCCCTATATTCATCAACCAAC 180
Db 5267 TTCTGGGCTGACCGGGTGTGATTTCTCAGCCCTTTCGCAATCCCTATATTCATCAACCAAC 5326
QY 181 CCCTTGGCGCGCGATGTACCGCTGCGCGCGCGCGCTGACCTCTGCTTTCGCAACCGCGC 240
Db 5327 CCCTTGGCGCGCGATGTACCGCTGCGCGCGCGCGCTGACCTCTGCTTTCGCAACCGCGC 5386
QY 241 CGACCACTCGGCTCGCTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGTTCAGTCTGT 300
Db 5387 CGACCACTCGGCTCGCTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGTTCAGTCTGT 5446
QY 301 AGACCTACACAGCTGGGCGCGCGCTTAACCGCGCTGCTCCGCGCGCGCGCGCGCGCG 360
Db 5447 AGACCTACACAGCTGGGCGCGCGCTTAACCGCGCTGCTCCGCGCGCGCGCGCGCGCG 5506
QY 361 CCAGTGCCTGATGTGCACTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Db 5507 CCAGTGCCTGATGTGCACTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5566
QY 421 TCTCCCTTACCTCTTCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 5567 TCTCCCTTACCTCTTCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5626
QY 481 AGTCCGCTTTTACCCCTTTCAGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 5627 AGTCCGCTTTTACCCCTTTCAGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5686
QY 541 AATATGCCAGTACCGGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 5687 AATATGCCAGTACCGGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5746
QY 601 GCTGTGCGGGTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 5747 GCTGTGCGGGTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5806
QY 661 TCCGTTGATATCAATTCAATAACCTTCGACCGGATGTTTCGTTATTTAGTCAGCGCGCGCG 720
Db 5807 TCCGTTGATATCAATTCAATAACCTTCGACCGGATGTTTCGTTATTTAGTCAGCGCGCGCG 5866
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Db 5867 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5926
```


QY	781	GTGAGACCTCTGGGGTGGCTGAGGAGGGCTACCTCTGGTCTGCTTATGCTTTGCATA	840	Db	7007	CATACCTTTTGATGATTTCTGCCAGAGTGCCGCCCTTGGCCTTCAGGGCTGCGCTTTC	7066
Db	5927	GTGAGACCTCTGGGGTGGCTGAGGAGGGCTACCTCTGGTCTGCTTATGCTTTGCATA	5986	QY	1921	CAGTCTACTGTCGTAGCTTACAGCCCTTAAGATGAAGTGGTAAACTCGGAGTTG	1980
QY	841	CATGGCTCACTGTAATTCCTATATACACCTTATACCGGTGCCCTCGGGCTGTTG	900	Db	7067	CAGTCTACTGTCGTAGCTTACAGCCCTTAAGATGAAGTGGTAAACTCGGAGTTG	7126
Db	5987	CATGGCTCACTGTAATTCCTATATACACCTTATACCGGTGCCCTCGGGCTGTTG	6046	QY	1981	TAG 1983	
QY	901	GACTTTGCCCTTGAGCTTGAGTTTCGAACTTTACCCCGGTAAACCAATACGCGGTC	960	Db	7127	TAG 7129	
Db	6047	GACTTTGCCCTTGAGCTTGAGTTTCGAACTTTACCCCGGTAAACCAATACGCGGTC	6106	RESULT 6			
QY	961	TCCCGTTATTCAGCACTGCTGCCACCGCTTCTGTCGGGTGCGGAGCTGCGAG	1020	ID	AAV54729	standard; DNA; 7195 BP.	
Db	6107	TCCCGTTATTCAGCACTGCTGCCACCGCTTCTGTCGGGTGCGGAGCTGCGAG	6166	XX	AAV54729;		
QY	1021	CTCACCACACGCGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT	1080	AC	25-MAR-2003 (revised)		
Db	6167	CTCACCACACGCGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT	6226	DT	30-OCT-1998 (first entry)		
QY	1081	GTCCGTGAGATCGCCCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT	1140	XX	DNA sequence of the Burmese isolate of ET-NANB.		
Db	6227	GTCCGTGAGATCGCCCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT	6286	XX	Enterically transmitted nonA/nonB hepatitis virus; identification; HEV;		
QY	1141	GGCGGCTCGCGACAGATTGATTTGTCGGCTGCTGGCCAGCTGTTCTACTCCCGTCCC	1200	KW	ET-NANB; detection; vaccine; ds.		
Db	6287	GGCGGCTCGCGACAGATTGATTTGTCGGCTGCTGGCCAGCTGTTCTACTCCCGTCCC	6346	XX	Hepatitis virus.		
QY	1201	GTTGTCTCAGCAATGGCGAGCGACTGTTAAAGTTGTATATCTGTAGAGAATGCTCAG	1260	Key	Location/Qualifiers		
Db	6347	GTTGTCTCAGCAATGGCGAGCGACTGTTAAAGTTGTATATCTGTAGAGAATGCTCAG	6406	FT	28..5109		
QY	1261	CAGGATAAGGGTATGCAATCCCGCATGACATTCGCGAGAACTCTGTGTTGTTATT	1320	FT	/*tag= a		
Db	6407	CAGGATAAGGGTATGCAATCCCGCATGACATTCGCGAGAACTCTGTGTTGTTATT	6466	FT	/note= "see AAW71209"		
QY	1321	CAGGATTATGTAACCAACATGACAGATCGCGGACGCTTCTCCAGCCCATCGCG	1380	FT	5106..5477		
Db	6467	CAGGATTATGTAACCAACATGACAGATCGCGGACGCTTCTCCAGCCCATCGCG	6526	FT	/*tag= c		
QY	1381	CTTTCTCTCTCTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT	1440	FT	/note= "see AAW71211"		
Db	6527	CTTTCTCTCTCTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT	6586	FT	5147..7129		
QY	1441	GACAGTCACTTATAGCTCTTCGATCGGCCAGTTTATGTTCTGACTCTGTGACCTTG	1500	FT	/*tag= b		
Db	6587	GACAGTCACTTATAGCTCTTCGACTGGCCAGTTTATGTTCTGACTCTGTGACCTTG	6646	FT	/note= "see AAW71210"		
QY	1501	GTTAATGTTGCGCGCGCGCGAGCGCTTGCCTGCTCGATTGCAAGGTCAACA	1560	PN	US5789559-A.		
Db	6647	GTTAATGTTGCGCGCGCGCGAGCGCTTGCCTGCTCGATTGCAAGGTCAACA	6706	XX	04-AUG-1998.		
QY	1561	CTTGACGGTCCGCCCTCTCCACCATCAGCAGTACTCGAAGACCTTCTTTGCTCGCG	1620	XX	25-JUL-1994; 94US-00279823.		
Db	6707	CTTGACGGTCCGCCCTCTCCACCATCAGCAGTACTCGAAGACCTTCTTTGCTCGCG	6766	XX	17-JUN-1988; 88US-00208997.		
QY	1621	CTCCGGGTAAAGCTCTCTTCTGGAGCGAGGCAACATTAAGCGGGTACCTTAAAT	1680	PR	11-APR-1989; 89US-00336672.		
Db	6767	CTCCGGGTAAAGCTCTCTTCTGGAGCGAGGCAACATTAAGCGGGTACCTTAAAT	6826	PR	19-JUN-1989; 89US-00367486.		
QY	1681	TATAACACCACTGTAGCGACCAACTGCTTGTTCAGAAATCGCGCGGACCGGGTCCGT	1740	PR	13-OCT-1989; 89US-00420921.		
Db	6827	TATAACACCACTGTAGCGACCAACTGCTTGTTCAGAAATCGCGCGGACCGGGTCCGT	6886	PR	05-APR-1990; 90US-00505888.		
QY	1741	ATTTCCTTACACCACTAGCCTGGGTGCTGGTCTCCGCTCTCCATTTCTGCGGTTCCGCTT	1800	PR	05-APR-1991; 91US-00681078.		
Db	6887	ATTTCCTTACACCACTAGCCTGGGTGCTGGTCTCCGCTCTCCATTTCTGCGGTTCCGCTT	6946	PA	(GENE-) GENELABS TECHNOLOGIES INC.		
QY	1801	TTAGCCCCCACTCTGCGCTAGCATGCTTTCAGATACCTTGGACTACCTTCCCGGCC	1860	PI	Bradley DW, Reyes GR, Krawczynski KZ, Tam A, Fry KE, Yarbough PO;		
Db	6947	TTAGCCCCCACTCTGCGCTAGCATGCTTTCAGATACCTTGGACTACCTTCCCGGCC	7006	DR	WPI: 1998-446186/38.		
QY	1861	CATACTTTTGATGATTTCTTGCCCCAGAGTGCCGCCCTTGGCTTTCAGGGCTGCGCTTTC	1920	DR	P-PSDB; AAW71209, AAW71210, AAW71211.		

Query Match									
Best Local Similarity 100.0%; Score 1983; DB 2; Length 7195;									
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGCGCCCTCGGCTATATTTTGTGTGCTCCTCATGTTTTCCTATGCTGCGCGGCA	60						
Db	5147	ATGCGCCCTCGGCTATATTTTGTGTGCTCCTCATGTTTTCCTATGCTGCGCGGCA	5206						
QY	61	CGCGCGGTCAGCGCTCGGCGCGCTGCTGGCGGCGCAGCGCGGTTCCCGCGGTGT	120						
Db	5207	CGCGCGGTCAGCGCTCGGCGCGCTGCTGGCGGCGCAGCGCGGTTCCCGCGGTGT	5266						
QY	121	TTCTGGGGTGACGGGTGTAATCTCAGCCCTTCGCAATCCCTATATTCATCCAAACAA	180						
Db	5267	TTCTGGGGTGACGGGTGTAATCTCAGCCCTTCGCAATCCCTATATTCATCCAAACAA	5326						
QY	181	CCCTTCGCGCCCGATGTCAACGCTCGGCGCGGGCTGACCTCGTGTTCGCAACCCGCC	240						
Db	5327	CCCTTCGCGCCCGATGTCAACGCTCGGCGCGGGCTGACCTCGTGTTCGCAACCCGCC	5386						
QY	241	CGACCACTCGGCTCGGCTGGGCTGACGAGCGCGCAGCGCGCGCTGCTCAGTGTGT	300						
Db	5387	CGACCACTCGGCTCGGCTGGGCTGACGAGCGCGCAGCGCGCGCTGCTCAGTGTGT	5446						
QY	301	AGACCTACACAGCTGGGGCGCGCGCTAAACGGGCTCGCTCCGCGCCCATGACACCCCG	360						
Db	5447	AGACCTACACAGCTGGGGCGCGCGCTAAACGGGCTCGCTCCGCGCCCATGACACCCCG	5506						
QY	361	CAAGTGCCTGATGTGACCTCCGCGCGGCGCAATCTTGGCGCGGAGTATTAACCTATCA	420						
Db	5507	CAAGTGCCTGATGTGACCTCCGCGCGGCGCAATCTTGGCGCGGAGTATTAACCTATCA	5566						
QY	421	TCTCCCTTACCTCTTCGCTGCGCACCAGCACTAACCTGGTCTTTATGCGCGCCCTCT	480						
Db	5567	TCTCCCTTACCTCTTCGCTGCGCACCAGCACTAACCTGGTCTTTATGCGCGCCCTCT	5626						
QY	481	AGTCGGCTTTTACCCCTTCAGGACGGCAACCAATACCATATATATGCGCACGAACTCT	540						
Db	5627	AGTCGGCTTTTACCCCTTCAGGACGGCAACCAATACCATATATATGCGCACGAACTCT	5686						
QY	541	AATATGCCAGTACGGGTTCGCGTCCGACCAATCCGTTACCGCGCGCTGGTCCCAAT	600						
Db	5687	AATATGCCAGTACGGGTTCGCGTCCGACCAATCCGTTACCGCGCGCTGGTCCCAAT	5746						
QY	601	GCTGTGCGGGTTACGCACTCTCCATCTCATCTGCGCACAGACCAACCAACCCGAGC	660						
Db	5747	GCTGTGCGGGTTACGCACTCTCCATCTCATCTGCGCACAGACCAACCAACCCGAGC	5806						
QY	661	TCCGTTGATATGAATTCATTAACCTCGACGGATGTTCTGTTATTTAGTCCAGCCCGCAT	720						
Db	5807	TCCGTTGATATGAATTCATTAACCTCGACGGATGTTCTGTTATTTAGTCCAGCCCGCAT	5866						
QY	721	GCCTCTGAGCTGTGATCCCAAGTGAGCGCTPACACTATCGPAACCAAGGCTGGCGCTCC	780						
Db	5867	GCCTCTGAGCTGTGATCCCAAGTGAGCGCTPACACTATCGPAACCAAGGCTGGCGCTCC	5926						
QY	781	GTCGAGACCTCTGGGTGCTGAGGAGGAGGCTACTCTGGTCTTGTATGCTTTGCATA	840						
Db	5927	GTCGAGACCTCTGGGTGCTGAGGAGGAGGCTACTCTGGTCTTGTATGCTTTGCATA	5986						
QY	841	CATGGCTCACTCGTAAATTCCTATACTATAACACCTATACCGGTGCGCTCGGGCTGTG	900						
Db	5987	CATGGCTCACTCGTAAATTCCTATACTATAACACCTATACCGGTGCGCTCGGGCTGTG	6046						
QY	901	GACTTTGCCCTTGAAGCTTGAAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTC	960						
Db	6047	GACTTTGCCCTTGAAGCTTGAAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTC	6106						
QY	961	TCCCGTTATTCAGACACTGCTCGCACCGCTTGTGCGGGTGGGACGGGACTGCCGAG	1020						
Db	6107	TCCCGTTATTCAGACACTGCTCGCACCGCTTGTGCGGGTGGGACGGGACTGCCGAG	6166						

RESULT 7
AAA99259
ID AAA99259 standard; DNA; 7195 Bp.
XX

QY	1021	CTCACACACACGGCTGTACTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGCT	1080
Db	6167	CTCACACACACGGCTGTACTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGCT	6226
QY	1081	GTGCGTGAGATCGCGCGGGGATAGCCCTCAACCTGTTCACCTTGTCTGACACTCTGCTTT	1140
Db	6227	GTGCGTGAGATCGCGCGGGGATAGCCCTCAACCTGTTCACCTTGTCTGACACTCTGCTTT	6286
QY	1141	GCGCGCTCGCGACAGAAATGATTTCTGCGGCTGGTGGCCAGCTGTTCTACTCCCGTCCC	1200
Db	6287	GCGCGCTCGCGACAGAAATGATTTCTGCGGCTGGTGGCCAGCTGTTCTACTCCCGTCCC	6346
QY	1201	GTGTCTCAGCCAAATGCGGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAACTCTAG	1260
Db	6347	GTGTCTCAGCCAAATGCGGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAACTCTAG	6406
QY	1261	CAGGATAAGGGTATTGCAATCCCGATGACATGTGACCTCGGAGAACTCTGCTGTGTTAT	1320
Db	6407	CAGGATAAGGGTATTGCAATCCCGATGACATGTGACCTCGGAGAACTCTGCTGTGTTAT	6466
QY	1321	CAGGATTATGATACCAACATGAAACAAGATCGGCGAGCCCTTCTCCAGCCCATCGCGC	1380
Db	6467	CAGGATTATGATACCAACATGAAACAAGATCGGCGAGCCCTTCTCCAGCCCATCGCGC	6526
QY	1381	CTTTCTCTGTCTCTGAGCTAAATGATGTCTTTGGCTCTCTCTCAACGCTGCCGAGTAT	1440
Db	6527	CTTTCTCTGTCTCTGAGCTAAATGATGTCTTTGGCTCTCTCTCAACGCTGCCGAGTAT	6586
QY	1441	GACCACTCACTTATGGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG	1500
Db	6587	GACCACTCACTTATGGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG	6646
QY	1501	GTTAATGTTGCGACCGGCGCGAGCCGCTTGCCCGGTGCTCGATGGACCAAGTCA	1560
Db	6647	GTTAATGTTGCGACCGGCGCGAGCCGCTTGCCCGGTGCTCGATGGACCAAGTCA	6706
QY	1561	CTTGACGGTTCGCGCCCTCTCCACCATCCAGACGACTCTCGAAGACCTTCTTGTCTCGCG	1620
Db	6707	CTTGACGGTTCGCGCCCTCTCCACCATCCAGACGACTCTCGAAGACCTTCTTGTCTCGCG	6766
QY	1621	CTCCCGGTAAAGCTCTCTTTCTGGAGGACAGCACAACTTAAAGCCGGGTACCTTTATAAT	1680
Db	6767	CTCCCGGTAAAGCTCTCTTTCTGGAGGACAGCACAACTTAAAGCCGGGTACCTTTATAAT	6826
QY	1681	TATACACCACTGTAGCGACCACTGCTTCTGAGAAATGCGCGGGGACCGGCTCGCT	1740
Db	6827	TATACACCACTGTAGCGACCACTGCTTCTGAGAAATGCGCGGGGACCGGCTCGCT	6886
QY	1741	ATTTCCACTTACACCACTAGCTGGGTGCTGCTCCCGTCTCCATTTCTGCGGTTGCGGTT	1800
Db	6887	ATTTCCACTTACACCACTAGCTGGGTGCTGCTCCCGTCTCCATTTCTGCGGTTGCGGTT	6946
QY	1801	TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAAGATACCTTGGACTACCTTGCCCGGCC	1860
Db	6947	TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAAGATACCTTGGACTACCTTGCCCGGCC	7006
QY	1861	CATACCTTTTGTATGTTCTGCGCGAGAGTCCGCCCCCTTGGCTTCAAGGCTGCGCTTC	1920
Db	7007	CATACCTTTTGTATGTTCTGCGCGAGAGTCCGCCCCCTTGGCTTCAAGGCTGCGCTTC	7066
QY	1921	CAGTCTACTGCTCGCTGAGCTTTCAGCGCTTAAAGTGAAGGTGGGTAAACCTCGGAGTTG	1980
Db	7067	CAGTCTACTGCTCGCTGAGCTTTCAGCGCTTAAAGTGAAGGTGGGTAAACCTCGGAGTTG	7126
QY	1981	TAG 1983	
Db	7127	TAG 7129	

Db 6287 GCGCGCTCGCCGACAGAAATTGATTTCGTGCGCTGGTGGCCAGCTGTTCTACTCCGTC 6346
QY 1201 GTTGTCTCAGCCAAATGCGAGCGAGCTGTTAAGTTGTATACATCTGTAGGAATGCTCAG 1260
Db 6347 GTTGTCTCAGCCAAATGCGAGCGAGCTGTTAAGTTGTATACATCTGTAGGAATGCTCAG 6406
QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTCGACCTCGGAGAAATCTCGTGTGGTTATT 1320
Db 6407 CAGGATAAGGGTATTGCAATCCCGCATGACATTCGACCTCGGAGAAATCTCGTGTGGTTATT 6466
QY 1321 CAGGATTATGATACCAACATGAACATGACATTCGACCTCGGAGAAATCTCGTGTGGTTATT 1380
Db 6467 CAGGATTATGATACCAACATGAACATGACATTCGACCTCGGAGAAATCTCGTGTGGTTATT 6526
QY 1381 CTTTCTCTGCTCTCGAGCTAATGATGCTTTGCTCTCTCCTCAGCGTCCGAGTAT 1440
Db 6527 CTTTCTCTGCTCTCGAGCTAATGATGCTTTGCTCTCTCCTCAGCGTCCGAGTAT 6586
QY 1441 GACCACTCACTTATGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACTTTG 1500
Db 6587 GACCACTCACTTATGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACTTTG 6646
QY 1501 GTTATGTTGCGACCGCGCGAGCGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db 6647 GTTATGTTGCGACCGCGCGAGCGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6706
QY 1561 CTTGACGCTGCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTGCTGCTGCG 1620
Db 6707 CTTGACGCTGCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTGCTGCTGCG 6766
QY 1621 CTCGCGGTAAAGTCTCTTTCTGGGAGGAGGACCACTAAAGCCGGGTACCTTATATAT 1680
Db 6767 CTCGCGGTAAAGTCTCTTTCTGGGAGGAGGACCACTAAAGCCGGGTACCTTATATAT 6826
QY 1681 TATACACACACTCTAGACCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db 6827 TATACACACACTCTAGACCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6886
QY 1741 ATTTCACCTTACACACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 6887 ATTTCACCTTACACACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6946
QY 1801 TTAGCCCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 6947 TTAGCCCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7006
QY 1861 CATATCTTTGATGATTTCTGCCAGAGTCCCGCCCTTGGCTTTCAGGGCTGCGCTTTC 1920
Db 7007 CATATCTTTGATGATTTCTGCCAGAGTCCCGCCCTTGGCTTTCAGGGCTGCGCTTTC 7066
QY 1921 CAGTCTACTGCTGCTGAGCTTCCAGCGCTTAAAGATGAAGTGGTAAACTCGGGAGTTG 1980
Db 7067 CAGTCTACTGCTGCTGAGCTTCCAGCGCTTAAAGATGAAGTGGTAAACTCGGGAGTTG 7126
QY 1981 TAG 1983
Db 7127 TAG 7129

RESULT 8

AAF83495

ID AAF83495 standard; DNA; 7195 BP.

XX

AC AAF83495;

XX

DT 23-JUL-2001 (first entry)

XX

DE ET-NANB viral DNA sequence from HEV-Burma strain.

XX

KW HEV; enterically transmitted nonA/nonB viral hepatitis agent; viral;

KW

ET-NANB; infection; vaccine; ds.

XX

OS Hepatitis E virus.

XX PH Location/Qualifiers
FT CDS 28..5109
FT /*tag= a
FT /product= "ORF1"
FT CDS 5106..5477
FT /*tag= c
FT /product= "ORF3"
FT CDS 5147..7129
FT /*tag= b
FT /product= "ORF2"
XX US6229005-B1.
XX 08-MAY-2001.
XX 03-AUG-1998; 98US-00128275.
XX 17-JUN-1988; 88US-00208997.
XX 11-APR-1989; 89US-00336672.
XX 16-JUN-1989; 89US-00367486.
XX 13-OCT-1989; 89US-00420921.
XX 05-APR-1990; 90US-00505888.
XX 05-APR-1991; 91US-00681078.
XX 25-JUL-1994; 94US-00279823.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX Reyes GR, Yarbrough PO, Bradley DW, Krawczynski KZ, Tam AM,
XX Fry KE;
XX WPI; 2001-342705/36.
XX P-PSDB; AAB62522, AAB62523, AAB62524.
XX New DNA sequences of enterically transmitted non-A/non-B (ET-NANB)
XX hepatitis viral agent, useful in diagnosing infection by an enterically
XX transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.
XX Disclosure; Col 49-56; 45pp; English.
XX The invention relates to an isolated DNA comprising the genome of an
XX enterically transmitted nonA/nonB (ET-NANB) viral hepatitis agent (also
XX referred as HEV). The DNA sequences or their fragments are useful in
XX preparing ET-NANB viral proteins and as probes for virus detection. These
XX are particularly useful in diagnosing infection by an enterically
XX transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.
XX The present sequence represents a ET-NANB viral DNA sequence from HEV-
XX Burma strain
SQ Sequence 7195 BP; 1221 A; 2312 C; 1868 G; 1794 T; 0 U; 0 Other;
Query Match 100.0%; Score 1983; DB 4; Length 7195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGCCCTATTTGTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
Db 5147 ATGCGCCCTCGCCCTATTTGTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5206
QY 61 CGCGCCGCTCAGCGCTCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCTTCCGGCGTGGT 120
Db 5207 CGCGCCGCTCAGCGCTCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCTTCCGGCGTGGT 5266
QY 121 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTTATTTATTCATCCAAACCAAC 180
Db 5267 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTTATTTATTCATCCAAACCAAC 5326
QY 181 CCCTTGCCTCCCGATGTCACCGCTGCGCGCGCGCTGCGACCTCGTGTGTTGCGCAACCGCGCC 240
Db 5327 CCCTTGCCTCCCGATGTCACCGCTGCGCGCGCGCTGCGACCTCGTGTGTTGCGCAACCGCGCC 5386
QY 241 CGACCACTCGGCTCGGCTTGGCTGACCAAGCCCGCGCGCGCTTGGCTTGGCTCGGCTCGT 300

Db	5387	CGACCACTCGGCTCCGCTTGGGCTGACAGGCCAGCGCCCGCGCTTGCCTCACGTCG	5446
QY	301	AGACCTTACCACAGCTGGGGCGCGCGCTAAACCGCGGTGCTCCGGCCCAATGACACCCCG	360
Db	5447	AGACCTTACCACAGCTGGGGCGCGCGCTAAACCGCGGTGCTCCGGCCCAATGACACCCCG	5506
QY	361	CGAGTCCTGATGTGACCTCCCGCGGGCGCACTTTGGCGGGAGTAACTATCAACA	420
Db	5507	CCAGTGCCTGATGTGACCTCCCGCGGGCGCACTTTGGCGGGAGTAACTATCAACA	5566
QY	421	TCCTCCCTTACCTCTTCCGTTGGCCACCGGCACTAACTGTGTTCTTTATGCGCGCCCTCT	480
Db	5567	TCCTCCCTTACCTCTTCCGTTGGCCACCGGCACTAACTGTGTTCTTTATGCGCGCCCTCT	5626
QY	481	AGTCGCGCTTTACCCCTTCAGACGGGCAACCAATACCAATATAATATGGCCAGGAGCTTCT	540
Db	5627	AGTCGCGCTTTACCCCTTCAGACGGGCAACCAATACCAATATAATATGGCCAGGAGCTTCT	5686
QY	541	AATTATGCCAGTACCGGTTGCGGTGCGGCACCAATCGTTACCGCGCTGTGCCCAAT	600
Db	5687	AATTATGCCAGTACCGGTTGCGGTGCGGCACCAATCGTTACCGCGCTGTGCCCAAT	5746
QY	601	GCTGTGCGCGGTTACGCGCATCTCCATCTCATTTCTGGCCACAGACCAACCCCGGACG	660
Db	5747	GCTGTGCGCGGTTACGCGCATCTCCATCTCATTTCTGGCCACAGACCAACCCCGGACG	5806
QY	661	TCCGTTGATATGAATCAATAAACCCTGACGAGTTCGTATTTTAGTCCAGCCGGGATTA	720
Db	5807	TCCGTTGATATGAATCAATAAACCCTGACGAGTTCGTATTTTAGTCCAGCCGGGATTA	5866
QY	721	GCTCTGAGCTTGTGATCCCAAGTGAGGCGCTACACTATGTAACCAAGCTGCGCTCC	780
Db	5867	GCTCTGAGCTTGTGATCCCAAGTGAGGCGCTACACTATGTAACCAAGCTGCGCTCC	5926
QY	781	GTCGAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCTGCTCTTGTATGCTTTGCATA	840
Db	5927	GTCGAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCTGCTCTTGTATGCTTTGCATA	5986
QY	841	CATGGCTCACTCGTAAATTCCTATATACACCTTATACCGGTGACCTGCGGCTGTTG	900
Db	5987	CATGGCTCACTCGTAAATTCCTATATACACCTTATACCGGTGACCTGCGGCTGTTG	6046
QY	901	GACTTTGCCCTTGAGCTTGATTTGCAACCTTACCCCGGTACACCAATACCGGCTC	960
Db	6047	GACTTTGCCCTTGAGCTTGATTTGCAACCTTACCCCGGTACACCAATACCGGCTC	6106
QY	961	TCCCGTTATTCAGCACTGCTCGGCACCGCCTTCGTGCGGTGCGGACGGGACTGCCGAG	1020
Db	6107	TCCCGTTATTCAGCACTGCTCGGCACCGCCTTCGTGCGGTGCGGACGGGACTGCCGAG	6166
QY	1021	CTCACCAACGGCTGTACCGCTTTTAGAGGACCTCTATTTTACTAGTACTAATGTT	1080
Db	6167	CTCACCAACGGCTGTACCGCTTTTAGAGGACCTCTATTTTACTAGTACTAATGTT	6226
QY	1081	GTCGCTGAGATCGCGCGGATAGCCTCACCTGTTTCAACCTTGCTGACACTCTGCTT	1140
Db	6227	GTCGCTGAGATCGCGCGGATAGCCTCACCTGTTTCAACCTTGCTGACACTCTGCTT	6286
QY	1141	GGCGGCTTCCGACAGAAATGATTTGCTGCGCTGGTGCCAGCTGTTCTACTCCCGTCCC	1200
Db	6287	GGCGGCTTCCGACAGAAATGATTTGCTGCGCTGGTGCCAGCTGTTCTACTCCCGTCCC	6346
QY	1201	GTTGCTCAGCAATGGCGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG	1260
Db	6347	GTTGCTCAGCAATGGCGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG	6406
QY	1261	CAGGATAAGGTTATTGCAATCCCGCATGACATTTGACCTCGGAGAACTCTCGTGTGTTATT	1320
Db	6407	CAGGATAAGGTTATTGCAATCCCGCATGACATTTGACCTCGGAGAACTCTCGTGTGTTATT	6466
QY	1321	CAGGATTATGATAACCAATGAAACAAGATGGCGGACGCGCTTCTCCAGCCCAATCGCGC	1380
Db	6467	CAGGATTATGATAACCAATGAAACAAGATGGCGGACGCGCTTCTCCAGCCCAATCGCGC	6526
QY	1381	CTTTCTCTGTCTTCGAGCTAATGATGTGCTTTGCTCTCTCTCACCGCTGCGGAGTAT	1440
Db	6527	CTTTCTCTGTCTTCGAGCTAATGATGTGCTTTGCTCTCTCTCACCGCTGCGGAGTAT	6586
QY	1441	GACCACTCCACTTATGGCTCTTCGACTGGCCCAAGTTTATGTTCTGACTCTGTGACCTTG	1500
Db	6587	GACCACTCCACTTATGGCTCTTCGACTGGCCCAAGTTTATGTTCTGACTCTGTGACCTTG	6646
QY	1501	GTTAATGTTGGACCGCGCGGAGCGGCTTGGCGGTGCTCGATTGGACCAAGGTCACA	1560
Db	6647	GTTAATGTTGGACCGCGCGGAGCGGCTTGGCGGTGCTCGATTGGACCAAGGTCACA	6706
QY	1561	CTTGACGGTGGCGCCCTCTCCACCATCCAGACGACTCTCGAAGACCTTCTTTGTCCTGCCG	1620
Db	6707	CTTGACGGTGGCGCCCTCTCCACCATCCAGACGACTCTCGAAGACCTTCTTTGTCCTGCCG	6766
QY	1621	CTCCGCGGTAAGCTCTCTTTCTGGGAGGCGAGGCAAACTAAAGCCGGGTACCTTATAAT	1680
Db	6767	CTCCGCGGTAAGCTCTCTTTCTGGGAGGCGAGGCAAACTAAAGCCGGGTACCTTATAAT	6826
QY	1681	TATAACCACTCTAGCGACCAACTGCTTGTGAGAAATGCGCGCGGACCGGTCGCT	1740
Db	6827	TATAACCACTCTAGCGACCAACTGCTTGTGAGAAATGCGCGCGGACCGGTCGCT	6886
QY	1741	ATTTCCACTTACACCACTAGCCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTGCGCGTT	1800
Db	6887	ATTTCCACTTACACCACTAGCCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTGCGCGTT	6946
QY	1801	TTAGCCCCCACTCTCGGCTAGCAATTCCTTGGAGTACCTTGGACTACCTGCGCGGCC	1860
Db	6947	TTAGCCCCCACTCTCGGCTAGCAATTCCTTGGAGTACCTTGGACTACCTGCGCGGCC	7006
QY	1861	CATACCTTTGATGATTTCTGCCAGAGTGGCGCCCTTCGCTTCAAGGCTGCGCTTC	1920
Db	7007	CATACCTTTGATGATTTCTGCCAGAGTGGCGCCCTTCGCTTCAAGGCTGCGCTTC	7066
QY	1921	CAGTCTACTCTGCTGAGCTTTCAGCGCTTAAAGTGAAGTGGGTAACTCGGAGTTG	1980
Db	7067	CAGTCTACTCTGCTGAGCTTTCAGCGCTTAAAGTGAAGTGGGTAACTCGGAGTTG	7126
QY	1981	TAG 1983	
Db	7127	TAG 7129	
RESULT 9			
ID	AAL50386 standard; DNA; 7195 BP.		
XX	AAL50386;		
AC	AAL50386;		
XX	08-NOV-2002 (first entry)		
DT	Hepatitis E virus (Burma strain) gene sequence.		
DE	HEV; enterically transmitted nonA/nonB hepatitis viral agent;		
XX	Burma strain; bile; ds; gene; ORF1; ORF2; ORF3.		
KW	Hepatitis E virus.		
XX	Hepatitis E virus.		
OS	Hepatitis E virus.		
XX	Hepatitis E virus.		
FH	Key		
FT	CDS		
FT	28..5109		
FT	/tag= a		
FT	/product= "Hepatitis E virus (Burma strain) ORF1 protein"		
FT	5106..5477		
FT	/tag= c		
FT	/product= "Hepatitis E virus (Burma strain) ORF3 protein"		
FT	5147..7129		
FT	/tag= b		
FT	/product= "Hepatitis E virus (Burma strain) ORF2 protein"		
XX			

US6379891-B1.
30-APR-2002.
19-APR-2000; 2000US-00553427.
17-JUN-1988; 88US-00208997.
11-APR-1989; 89US-00336672.
16-JUN-1989; 89US-00367486.
13-OCT-1989; 89US-00420921.
05-APR-1990; 90US-00505888.
15-APR-1991; 91US-00681078.
25-JUL-1994; 94US-00279823.
07-JUN-1995; 95US-00478507.
(USSH) US DEPT HEALTH & HUMAN SERVICES.
(GENE-) GENELABS TECHNOLOGIES INC.
Reyes GR, Yarbrough PO, Bradley DW, Krawczynski KZ, Tam A, Fry KE;
WPI; 2002-517277/55.
P-PSDB; AA015698, AA015699, AA015700.
Detecting the presence of enterically transmitted nonA/nonB hepatitis
viral (HEV) agents in bile samples from infected humans and monkeys using
polymerase chain reaction.
Disclosure; Col 13-24; 61pp; English.
The invention comprises a method for detecting the presence of
enterically transmitted nonA/nonB hepatitis viral (HEV) agents in a
sample and isolating HEV agents or nucleic acid fragments produced by the
agent. The method utilises PCR - using bile from a human or cynomolgus
monkey actively infected with HEV as a source of the agent. The method of
the invention is used for detecting the presence of a viral agent in a
sample of cultured cells infected with the agent and isolating
enterically transmitted nonA/nonB HEV agents or nucleic acid fragments
produced by the agent. The present DNA sequence was isolated from a
Burmese strain of the Hepatitis E virus, the present DNA sequence has
three open reading frames ORF1, ORF2 and ORF3
Sequence 7195 BP; 1221 A; 2312 C; 1868 G; 1794 T; 0 U; 0 Other;
Query Match
Best Local Similarity 100.0%; Score 1983; DB 6; Length 7195;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCCTATTTTGTGCTGCTCCTCATGTTTGGCTATGCTGCGCGCCA 5206
DB 5147 ATGCGCCCTCGGCCTATTTTGTGCTGCTCCTCATGTTTGGCTATGCTGCGCGCCA 5206
QY 61 CCGCCGGTACCGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCTTCCGCGGTGT 120
DB 5207 CCGCCGGTACCGCTGCGCGCGCTGCTGCGCGCGCGCGCGCTTCCGCGGTGT 5266
QY 121 TTCTGGGGTGACGGGGTGATTTCTCAGCCCTTCGCAATCCCTATATTATCAACCAAC 180
DB 5267 TTCTGGGGTGACGGGGTGATTTCTCAGCCCTTCGCAATCCCTATATTATCAACCAAC 5326
QY 181 CCCTTTCGCCCCGATGTACCGCTGCGCGCGGGGTGACCTCGTGTTCGCAACCCCGCC 240
DB 5327 CCCTTTCGCCCCGATGTACCGCTGCGCGCGGGGTGACCTCGTGTTCGCAACCCCGCC 5386
QY 241 CGACCACTCGGCTCGCTTGGCGTGACAGGCCCGCGCGCGCGCTTCCGCTCAGTGT 300
DB 5387 CGACCACTCGGCTCGCTTGGCGTGACAGGCCCGCGCGCGCGCTTCCGCTCAGTGT 5446
QY 301 AGACCTACCAAGCTGGGGCGCGCGCTAACCGCGGTGCGCTCCGCGCCCATGACACCCCG 360
DB 5447 AGACCTACCAAGCTGGGGCGCGCGCTAACCGCGGTGCGCTCCGCGCCCATGACACCCCG 5506
QY 361 CCAGTGCCTGTATGTGACTCCGCGCGCGCAATCTTTCGCGCGCGAGTATAACCTATCAACA 420
DB |||||

DB 5507 CCAGTGCCTGTATGTGACTCCGCGCGCGCAATCTTGC CGCGCGAGTATAACCTATCAACA 5566
QY 421 TCTCCCCCTTACCTCTTCCTGCGCCACCGGCACTAACTGCTGTTCTTTATGCGCGCCCTCTT 480
DB 5567 TCTCCCCCTTACCTCTTCCTGCGCCACCGGCACTAACTGCTGTTCTTTATGCGCGCCCTCTT 5626
QY 481 AGTCCGCTTTTACCCCTTCAGAGCGGCAACCAATACCCATATATATGCGCAAGGCTTCT 540
DB 5627 AGTCCGCTTTTACCCCTTCAGAGCGGCAACCAATACCCATATATATGCGCAAGGCTTCT 5686
QY 541 AATTATGCCAGTACCGGGTTGCCCGTGCCACAATCCGTTACCGCGCGCTGCTGCCCAAT 600
DB 5687 AATTATGCCAGTACCGGGTTGCCCGTGCCACAATCCGTTACCGCGCGCTGCTGCCCAAT 5746
QY 601 GCTGTCCGCGGTACCGCATCTCCATCTCATTTCTTGCCACAGACACACACCCCGGAG 660
DB 5747 GCTGTCCGCGGTACCGCATCTCCATCTCATTTCTTGCCACAGACACACACCCCGGAG 5806
QY 661 TCCGTTGATATGAATTCATAATAACCTCGAGGATGTTGTTATTTAGTCCAGCCCGGCATA 720
DB 5807 TCCGTTGATATGAATTCATAATAACCTCGAGGATGTTGTTATTTAGTCCAGCCCGGCATA 5866
QY 721 GCCTCTGAGCTTGTGATCCCAAGTACGCGCTACACTATCGTAACCAAGGCTGCGCTCC 780
DB 5867 GCCTCTGAGCTTGTGATCCCAAGTACGCGCTACACTATCGTAACCAAGGCTGCGCTCC 5926
QY 781 GTCGAGACCTCTGGGTGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTCGATA 840
DB 5927 GTCGAGACCTCTGGGTGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTCGATA 5986
QY 841 CATGGCTACTCGTAAATTCCTATATAACCTATACCTATACCGTGCCTCGGCGCTGTTG 900
DB 5987 CATGGCTACTCGTAAATTCCTATATAACCTATACCGTGCCTCGGCGCTGTTG 6046
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGTC 960
DB 6047 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGTC 6106
QY 961 TCCGCTTATTCAGACACTGCTCGCACCGCTTGTGCGCGGTGCGGACGCGACTGCGGAG 1020
DB 6107 TCCCGTTATTCAGACACTGCTCGCACCGCTTGTGCGCGGTGCGGACGCGACTGCGGAG 6166
QY 1021 CTCACACACCGCTGTACCCGCTTTATGAGGAGCTCTATTCTACTAGTACTATAGT 1080
DB 6167 CTCACACACCGCTGTACCCGCTTTATGAGGAGCTCTATTCTACTAGTACTATAGT 6226
QY 1081 GTCGCTGAGATCGGCGCGGGATAGCCCTCAACCTGTTCAACCTTGCTGACACTCTGCTT 1140
DB 6227 GTCGCTGAGATCGGCGCGGGATAGCCCTCAACCTGTTCAACCTTGCTGACACTCTGCTT 6286
QY 1141 GCGGCGCTGCGGACAGAAATGATTTGCTGCGCTGGTGGCGAGCTGTTCTACTCCCGTCC 1200
DB 6287 GCGGCGCTGCGGACAGAAATGATTTGCTGCGCTGGTGGCGAGCTGTTCTACTCCCGTCC 6346
QY 1201 GTTGTCTCAGCCAAATGGCGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
DB 6347 GTTGTCTCAGCCAAATGGCGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 6406
QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT 1320
DB 6407 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT 6466
QY 1321 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT 1380
DB 6467 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT 6526
QY 1381 CTTTTCTCTCTCGAGCTAAATGATGCTGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
DB 6527 CTTTTCTCTCTCGAGCTAAATGATGCTGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 6586
QY 1441 GACCACTCCACTTATGCTCTTCGACTGCGCGCGCAATCTTTCGCTGACCTGACCTTATCAACA 1500
DB 6587 GACCACTCCACTTATGCTCTTCGACTGCGCGCGCAATCTTTCGCTGACCTGACCTTATCAACA 6646

QY 661 TCCTGTTGATGATTAATCAATAAACCCTCGACGGATGTTGCTATTTTAGTCCAGCCCGGCATA 720
Db 5807 TCCTGTTGATGATTAATCAATAAACCCTCGACGGATGTTGCTATTTTAGTCCAGCCCGGCATA 5866
QY 721 GCCTCTGAGCTTGTGATCCCAAGTAGCGGCTACATATCGTAACCAAGGCTCGCGCTCC 780
Db 5867 GCCTCTGAGCTTGTGATCCCAAGTAGCGGCTACATATCGTAACCAAGGCTCGCGCTCC 5926
QY 781 GTGAGACCTCTGGGCTGGCTGAGGAGGAGGCTACCTCTGCTTGTATGCTTTGCATTA 840
Db 5927 GTGAGACCTCTGGGCTGGCTGAGGAGGAGGCTACCTCTGCTTGTATGCTTTGCATTA 5986
QY 841 CATGGCTCACTCGCTAAATTCCTATACCTAATACACCCCTATACCGGTGCGCTCGGGCTGTTG 900
Db 5987 CATGGCTCACTCGCTAAATTCCTATACCTAATACACCCCTATACCGGTGCGCTCGGGCTGTTG 6046
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTTAACACCAATACCGGGTC 960
Db 6047 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTTAACACCAATACCGGGTC 6106
QY 961 TCCCGTTATCCAGCACTGCTCGCACCGCTTCGTGGGTCGGAGCGGACTGCGGAG 1020
Db 6107 TCCCGTTATCCAGCACTGCTCGCACCGCTTCGTGGGTCGGAGCGGACTGCGGAG 6166
QY 1021 CTCACCAACACCGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGTT 1080
Db 6167 CTCACCAACACCGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGTT 6226
QY 1081 GTCGGTGAGATCGGCGCGGATAGCCCTACCCCTGTTCAACCTTCTGACACTCTGCTT 1140
Db 6227 GTCGGTGAGATCGGCGCGGATAGCCCTACCCCTGTTCAACCTTCTGACACTCTGCTT 6286
QY 1141 GCGGCTCGCGACAGAAATGATTTGCTGGCTGGTGGCAGCTGTTCTACTCCGTCOC 1200
Db 6287 GCGGCTCGCGACAGAAATGATTTGCTGGCTGGTGGCAGCTGTTCTACTCCGTCOC 6346
QY 1201 GTTGTCTCAGCCAAATGGCGAGCGACTGTTAAGTGTATATACATCTGTAGAGAAATGCTCAG 1260
Db 6347 GTTGTCTCAGCCAAATGGCGAGCGACTGTTAAGTGTATATACATCTGTAGAGAAATGCTCAG 6406
QY 1261 CAGGATAAGGGTATTGCAATTCGCGCATGACATTCGAGAAATCTCGTGTGTTATT 1320
Db 6407 CAGGATAAGGGTATTGCAATTCGCGCATGACATTCGAGAAATCTCGTGTGTTATT 6466
QY 1321 CAGGATTATGATAACCAATGAACATGACATCGCGAGCGCTTCTCCAGCCCATCGCGC 1380
Db 6467 CAGGATTATGATAACCAATGAACATGACATCGCGAGCGCTTCTCCAGCCCATCGCGC 6526
QY 1381 CCTTTCTCTGCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCCGAGTAT 1440
Db 6527 CCTTTCTCTGCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCCGAGTAT 6586
QY 1441 GACCAGTCACATATGGCTCTTCGACTGCCCGCAGTTTATGTTCTGACTCTGTGACCTTG 1500
Db 6587 GACCAGTCACATATGGCTCTTCGACTGCCCGCAGTTTATGTTCTGACTCTGTGACCTTG 6646
QY 1501 GTTAATGTTGCGACCGCGCGAGCGCTTGCCCGCTCGCTCGATTGGACCAAGTCA 1560
Db 6647 GTTAATGTTGCGACCGCGCGAGCGCTTGCCCGCTCGCTCGATTGGACCAAGTCA 6706
QY 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCGCG 1620
Db 6707 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCGCG 6766
QY 1621 CTCGCGGTAAGCTCTCTTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCTTATAT 1680
Db 6767 CTCGCGGTAAGCTCTCTTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCTTATAT 6826
QY 1681 TATAACACCTGCTAGACCACTGCTTTGCGAGAAATGCGCGCGGCAACCGGTCGCT 1740
Db 6827 TATAACACCTGCTAGACCACTGCTTTGCGAGAAATGCGCGCGGCAACCGGTCGCT 6886

QY 1741 ATTTCCACTTACACCACCTAGCCTGGGTGCTGGTCCCGTCTCCATTTCTGGGTTGCGTT 1800
Db 6887 ATTTCCACTTACACCACCTAGCCTGGGTGCTGGTCCCGTCTCCATTTCTGGGTTGCGTT 6946
QY 1801 TTAGCCCCCACCCTCTCGCTAGCAITTCCTTTAGGATACCTTTGGACTACCTTGCCTGCCGCC 1860
Db 6947 TTAGCCCCCACCCTCTCGCTAGCAITTCCTTTAGGATACCTTTGGACTACCTTGCCTGCCGCC 7006
QY 1861 CATACCTTTGATGATTTCTGCCAGAGTGGCGCCCTTGGCTTCAGGCTGGGCTTTC 1920
Db 7007 CATACCTTTGATGATTTCTGCCAGAGTGGCGCCCTTGGCTTCAGGCTGGGCTTTC 7066
QY 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAAGATGAAGTGGTAAACTCGGGAGTTG 1980
Db 7067 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAAGATGAAGTGGTAAACTCGGGAGTTG 7126
QY 1981 TAG 1983
Db 7127 TAG 7129

RESULT 11
AAQ14412
ID AAQ14412 standard; DNA; 7195 BP.
XX
AC AAQ14412;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 16-JAN-1992 (first entry)
XX
DE Forward strand of Burmese ET-NANB viral strain.
XX
KW Enterically transmitted non-A, non-B hepatitis virus; hepatitis C; HCV;
E.coli strain BB4; ATCC deposit number 67717; Burma; ss.
XX
OS Non-A.
non-B hepatitis virus.
XX
FH Key Location/Qualifiers
CDS 28..5109
FT /*tag= a
FT /note= "ORF 1"
CDS 5106..5477
FT /*tag= c
FT /note= "ORF 3"
CDS 5147..7129
FT /*tag= b
FT /note= "ORF 2"
XX
WO9115603-A.
PN
XX
PD 17-OCT-1991.
XX
PF 05-APR-1990; 90US-00505888.
XX
PR 05-APR-1990; 90US-00505888.
XX
PA (GENE-) GENELABS INC.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PA (USSH) US DEPT OF HEALTH.
XX
PI Reyes GR, Yarbrough PO, Bradley DW, Krawczynsk KZ, Tam A, Fry KB;
XX
DR WPI; 1991-325242/44.
DR P-PSDB; AAR14618, AAR14619, AAR14620.
XX
PT New viral proteins from non A-non-B hepatitis agent - used to treat and
prevent enterically-transmitted non-A non-B hepatitis virus.
XX
PS Claim 13; Page 97; 117pp; English.
XX
CC A positive clone ET1.1 was identified in a library prepared from bile of

CC cynomolgus monkeys infected with the Burma strain of ET-NANB. Both
CC strands of Etl.1 were sequenced. Identity of the sequence with sequences
CC in etiologic agents has been confirmed by locating a similar sequence in
CC a viral strain isolated in Burma. The sequence of the Burma strain is
CC given here. See AA014410 for Etl.1. (Updated on 25-MAR-2003 to correct PA
CC field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 7195 BP; 1221 A; 2309 C; 1859 G; 1806 T; 0 U; 0 Other;

Query Match 99.4%; Score 1970.2; DB 2; Length 7195;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1975; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy	1	ATGGGCGCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGCTGCGCGGCCA	60
Db	5147	ATGGGCGCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGCTGCGCGGCCA	5206
Qy	61	CCGCGCGGTGAGCGCTGCGCGCGGTGCTGCGCGCGCGCGCGGTTCGGGCGGTGGT	120
Db	5207	CCGCGCGGTGAGCGCTGCGCGCGGTGCTGCGCGCGCGCGGTTCGGGCGGTGGT	5266
Qy	121	TTCTGGGGTGACCGGGTGATTTCTCAGCCCTTGCAGTCCCTATATATCCAAACCAAC	180
Db	5267	TTCTGGGGTGACCGGGTGATTTCTCAGCCCTTGCAGTCCCTATATATCCAAACCAAC	5326
Qy	181	CCCTTGGCGCGCGATGTCAGCGTTCGCGCGCGGTGCGCGCGGTGCGCGCGGTGCG	240
Db	5327	CCCTTGGCGCGCGATGTCAGCGTTCGCGCGCGGTGCGCGCGGTGCGCGCGGTGCG	5386
Qy	241	CGACCACTCGGCTCCGCTTGGCGGTGACGAGCGCGCGCGCGCGGTGCGCTCAGCTCGT	300
Db	5387	CGACCACTCGGCTCCGCTTGGCGGTGACGAGCGCGCGCGCGGTGCGCTCAGCTCGT	5446
Qy	301	AGACCTACCAAGCTGGGGCGCGCGCTAACCGCGGTGCGCGCGCGGTGCGCGCGGTG	360
Db	5447	AGACCTACCAAGCTGGGGCGCGCGCTAACCGCGGTGCGCGCGCGGTGCGCGCGGTG	5506
Qy	361	CCAGTGCCTGATGTCGAGTCCGCGCGCGCGCGCGGTGCGCGCGGTGCGCGCGGTG	420
Db	5507	CCAGTGCCTGATGTCGAGTCCGCGCGCGCGCGCGGTGCGCGCGGTGCGCGCGGTG	5566
Qy	421	TCTCCCTTACCTCTCCGCTGGCGCACCGGCACTAACCTGGTTCCTTATGCGCGCGCTCT	480
Db	5567	TCTCCCTTACCTCTCCGCTGGCGCACCGGCACTAACCTGGTTCCTTATGCGCGCGCTCT	5626
Qy	481	AGTCCGCTTTTACCTCTCAGGACGCGCACTAACCTAACCTAACCTAACCTAACCT	540
Db	5627	AGTCCGCTTTTACCTCTCAGGACGCGCACTAACCTAACCTAACCTAACCTAACCT	5686
Qy	541	AATATGCCAGTACCGGGTTCGCGTGCACCAATCCGTTACCGCGCTGCTCCCAAT	600
Db	5687	AATATGCCAGTACCGGGTTCGCGTGCACCAATCCGTTACCGCGCTGCTCCCAAT	5746
Qy	601	GCTGTCCGCGGTATACCGCATCTCCATCTCATTTCTGGCCACAGACCAACCCCGGACG	660
Db	5747	GCTGTCCGCGGTATACCGCATCTCCATCTCATTTCTGGCCACAGACCAACCCCGGACG	5806
Qy	661	TCCGTTGATATGAATTCATTAACCTCGACGAGTGTTCGATTTAGTCCAGCGCGCAT	720
Db	5807	TCCGTTGATATGAATTCATTAACCTCGACGAGTGTTCGATTTAGTCCAGCGCGCAT	5866
Qy	721	GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGCGCTCC	780
Db	5867	GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGCGCTCC	5926
Qy	781	GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGCTTGTGTTATGCTTTGCA	840
Db	5927	GTCGAGACCTCTGCTGGTGGCTGAGGAGGAGGCTACCTCTGCTTGTGTTATGCTTTGCA	5986
Qy	841	CATGGCTCACTCGTAAATTCCTATACCTATACCTATACCTATACCTATACCTATG	900
Db	5987	CATGGCTCACTCGTAAATTCCTATACCTATACCTATACCTATACCTATACCTATG	6046

Qy	901	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAACCAATACCGGGTTC	960
Db	6047	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAACCAATACCGGGTTC	6106
Qy	961	TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTCGCGGTGGGACGAGTCCGCGAG	1020
Db	6107	TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTCGCGGTGGGACGAGTCCGCGAG	6166
Qy	1021	CTCACCACCAAGGCTGCTACCCGCTTTATGAAGAGCTCTATTTTACTAGTACTAATGGT	1080
Db	6167	CTCACCACCAAGGCTGCTACCCGCTTTATGAAGAGCTCTATTTTACTAGTACTAATGGT	6226
Qy	1081	GTCCGTTGAGATCGCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTCTGCTT	1140
Db	6227	GTCCGTTGAGATCGCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTCTGCTT	6286
Qy	1141	GGCGGCTGCGGACAGAAATTTGTTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1200
Db	6287	GGCGTCTCGCGGACAGAAATTTGTTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	6346
Qy	1201	GTGTCCTCAGCAATGCGGAGCGGCTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG	1260
Db	6347	GTGTCCTCAGCAATGCGGAGCGGCTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG	6406
Qy	1261	CAGGATAAGGGTATTTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGGTTATT	1320
Db	6407	CAGGATAAGGGTATTTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGGTTATT	6466
Qy	1321	CAGGATATGATTAACCAATGAACAGATCGGCGGAGCGGCTTCTCAGCGCCATCGCGC	1380
Db	6467	CAGGATATGATTAACCAATGAACAGATCGGCGGAGCGGCTTCTCAGCGCCATCGCGC	6526
Qy	1381	CCCTTCT	1440
Db	6527	CCCTTCT	6586
Qy	1441	GACCACTTCACTTATGCTCTCTGACCTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG	1500
Db	6587	GACCACTTCACTTATGCTCTCTGACCTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG	6646
Qy	1501	GTTAATGTTGGGACCGCGCGCGCGCTTTCGCGGTGCGTCTGATTCGACCAAGGTGACA	1560
Db	6647	GTTAATGTTGGGACCGCGCGCGCGCTTTCGCGGTGCGTCTGATTCGACCAAGGTGACA	6706
Qy	1561	CTTGACGCTCGCCCTCTCCACCATCCAGCACTCTGAGAGCTTCTTGTCTGCTGCGG	1620
Db	6707	CTTGACGCTCGCCCTCTCTCCACCATCCAGCACTCTGAGAGCTTCTTGTCTGCTGCGG	6766
Qy	1621	CTCCGCGGTAAAGCTCTCTTTCTGGGAGGCGAGGCAAACTAAAGCGGGTACCTTTATAT	1680
Db	6767	CTCCGCGGTAAAGCTCTCTTTCTGGGAGGCGAGGCAAACTAAAGCGGGTACCTTTATAT	6826
Qy	1681	TATAACACCACTGCTAGGACCAATCTGTTGAGAAATGCGCGCGGCGGCGGCGGCTGCT	1740
Db	6827	TATAACACCACTGCTAGGACCAATCTGTTGAGAAATGCGCGCGGCGGCGGCGGCTGCT	6886
Qy	1741	ATTTTCCACTTACCACTAGCTGCGGTGCTGCTCCCTCTCCATTTCTGCGGTTGCGGCTT	1800
Db	6887	ATTTTCCACTTACCACTAGCTGCGGTGCTGCTCCCTCTCCATTTCTGCGGTTGCGGCTT	6946
Qy	1801	TTAGCCCCCACTCTGCGCTAGCAATGCTTGAAGATACCTTGAAGTACCTTCCCTCCCGGCG	1860
Db	6947	TTAGCCCCCACTCTGCGCTAGCAATGCTTGAAGATACCTTGAAGTACCTTCCCTCCCGGCG	7006
Qy	1861	CATATTTTGTATGATTTCTGCCAGAGTGGCGGCGGCTTGGCTTCTGAGGCTGCGCTTTC	1920
Db	7007	CATATTTTGTATGATTTCTGCCAGAGTGGCGGCGGCTTGGCTTCTGAGGCTGCGCTTTC	7066
Qy	1921	CAGTCTACTGTCGTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTGAGTAAACTCGGAGTTG	1980
Db	7067	CAGTCTACTGTCGTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTGAGTAAACTCGGAGTTG	7126

QY 1981 TAG 1983
 Db 7127 TAG 7129

RESULT 12
 AAV61688
 ID AAV61688 standard; DNA; 1983 BP.
 XX AAV61688;
 XX 03-DEC-1998 (first entry)
 XX Hepatitis E virus hollow particle protein DNA #2.
 XX Hollow particle protein; virus; antibody; detection; immunoassay;
 KW infection; ss.
 XX Hepatitis virus.
 OS
 XX JP10234383-A.
 PN 08-SEP-1998.
 PD 28-FEB-1997; 97JP-00062445.
 XX 28-FEB-1997; 97JP-00062445.
 XX (DENK-) DENKA SEIKEN KK.
 PA (KOKU-) KOKURITSU YOB0 EISEI KENKYUSHO.
 XX WPI; 1998-535037/46.
 DR P-PSDB; AAW76369.
 XX
 PT Hepatitis E virus hollow particle poly:peptide(s) and nucleic acids
 PT encoding it - useful for more accurate detection of HEV in samples, using
 PT immuno-assays and nucleic acid hybridisation.
 XX
 PS Claim 13; Page 24-26; 29pp; Japanese.
 XX
 CC This sequence encodes a Hepatitis E viral hollow particle protein. This
 CC polypeptide can be used to raise antibodies to detect HEV infection in
 CC samples, e.g. by immuno-assay based techniques, and the nucleic acid can
 CC be used for the same in nucleic acid hybridisation assays. The
 CC polypeptides and nucleic acids allow more accurate detection of HEV than
 CC previously possible
 XX
 SQ Sequence 1983 BP; 342 A; 665 C; 485 G; 491 T; 0 U; 0 Other;

Query Match 98.3%; Score 1949.4; DB 2; Length 1983;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1962; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCGCCCTATTTTGTGTCCTCCTCAGTTTTCCTATGTCGCGCGGCA 60
 Db 1 ATGCGCCCTCGCCCTATTTTGTGTCCTCCTCAGTTTTCCTATGTCGCGCGGCA 60
 QY 61 CGCGCCGCTCAGCGCTGCGCGCGCGCTGTCGCGCGCGCGCGCGCGCGCTGCT 120
 Db 61 CGCGCCGCTCAGCGCTGCGCGCGCGCTGTCGCGCGCGCGCGCGCGCGCTGCT 120
 QY 121 TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTTCGCAATCCCTATATTCACCAACAC 180
 Db 121 TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTTCGCAATCCCTATATTCACCAACAC 180
 QY 181 CCTTCGCCCCCGATGTACACCGCTGCGCGCGCGCGCGCGCGCTGTCGCGCAACCGGCC 240
 Db 181 CCTTCGCCCCCGATGTACACCGCTGCGCGCGCGCGCGCGCGCTGTCGCGCAACCGGCC 240
 QY 241 CGACCACTCGGCTCCGCTTGGCGTGNACAGGCGCGCGCGCGCGCGCTTGCCTCAGTCGT 300
 Db 241 CGACCACTCGGCTCCGCTTGGCGTGNACAGGCGCGCGCGCGCGCGCTTGCCTCAGTCGT 300

QY 301 AGACCTACCAAGCTGGGGCGCGCGCTAAACCGCGTGGCTCGCTCCGCCCATGACACCCCG 360
 Db 301 AGACCTACCAAGCTGGGGCGCGCGCTAAACCGCGTGGCTCGCTCCGCCCATGACACCCCG 360
 QY 361 CCAGTGCCTGATGTGACTCCCGCGGCGCAATCTTTCGCGCGGAGTATTAACCTATCAACA 420
 Db 361 CCAGTGCCTGATGTGACTCCCGCGGCGCAATCTTTCGCGCGGAGTATTAACCTATCAACA 420
 QY 421 TCTCCCTTACCTCTTCGCTGGCGCACCGGCACTTAACCTGGTCTTTTATGCGCGCTCTT 480
 Db 421 TCTCCCTTACCTCTTCGCTGGCGCACCGGCACTTAACCTGGTCTTTTATGCGCGCTCTT 480
 QY 481 AGTCGCTTTTACCTCTTCAGGACGCAACCAATACCCATATATATGACCAAGCTTCT 540
 Db 481 AGTCGCTTTTACCTCTTCAGGACGCAACCAATACCCATATATATGACCAAGCTTCT 540
 QY 541 AATTATGCCAGTACCGGGTTGCGGTGCGCAATCCGTTACCGCGCTGCTGCCAAT 600
 Db 541 AATTATGCCAGTACCGGGTTGCGGTGCGCAATCCGTTACCGCGCTGCTGCCAAT 600
 QY 601 GCTGTCGGGCTTACGCCATCTCCATCTCATTTCTGSCACAGACCAACACCCCGGAG 660
 Db 601 GCTGTCGGGCTTACGCCATCTCCATCTCATTTCTGSCACAGACCAACACCCCGGAG 660
 QY 661 TCCGTTGATGATTAATTAACCTCGACGATGTTCTGTTATTTAGTCCAGCGCCGCGATA 720
 Db 661 TCCGTTGATGATTAATTAACCTCGACGATGTTCTGTTATTTAGTCCAGCGCCGCGATA 720
 QY 721 GCCTCTGAGCTTGTGATCCAAAGTGAAGGCTTACACTATCGTAAACAAAGCTGGGCTCC 780
 Db 721 GCCTCTGAGCTTGTGATCCAAAGTGAAGGCTTACACTATCGTAAACAAAGCTGGGCTCC 780
 QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACTCTGCTGTTGTTATGCTTGCATA 840
 Db 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACTCTGCTGTTGTTATGCTTGCATA 840
 QY 841 CATGGCTCACTCGTAAATTCCTATATAACACCTTACCGGTCCTCCGGGCTGTTG 900
 Db 841 CATGGCTCACTCGTAAATTCCTATATAACACCTTACCGGTCCTCCGGGCTGTTG 900
 QY 901 GACTTGGCTTGAAGTGTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTC 960
 Db 901 GACTTGGCTTGAAGTGTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTC 960
 QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCTGCGGTGCGGAGGAGTCCCGAG 1020
 Db 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCTGCGGTGCGGAGGAGTCCCGAG 1020
 QY 1021 CTCACACCAACCGCTGCTACCGCTTTTATGAAGGACCTCTATTTTACTAGTACTAATGCT 1080
 Db 1021 CTCACACCAACCGCTGCTACCGCTTTTATGAAGGACCTCTATTTTACTAGTACTAATGCT 1080
 QY 1081 GTCGCTGAGATCGCGCGGGATAGCCCTCACCCTGTTCAACCTTGTGACACTCTGCTT 1140
 Db 1081 GTCGCTGAGATCGCGCGGGATAGCCCTCACCCTGTTTAACTTGTGACACTCTGCTT 1140
 QY 1141 GCGCGCTTCCGACAGAAATGATTTTCGTCGGCTGGTGGCCAGCTGTTCTACTCCCGTCC 1200
 Db 1141 GCGCGCTTCCGACAGAAATGATTTTCGTCGGCTGGTGGCCAGCTGTTCTACTCCCGTCC 1200
 QY 1201 GTTGTCTCAGCAATATGGCGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAACTCTAG 1260
 Db 1201 GTTGTCTCAGCAATATGGCGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAACTCTAG 1260
 QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTCGCGGAGAACTCTCGTGTGTTATT 1320
 Db 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTCGCGGAGAACTCTCGTGTGTTATT 1320
 QY 1321 CAGGATTATGATAACCAATGACAAAGATCGCGCGAGCGCTTCTTCAGGCCCATCGCGC 1380
 Db 1321 CAGGATTATGATAACCAATGACAAAGATCGCGCGAGCGCTTCTTCAGGCCCATCGCGC 1380
 QY 1381 CCTTCTCTGCTCTTCGAGCTAATGATGTGCTTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440

QY 841 CATGGCTCACTCGTAATTCCTATATACTAATACACCCCTATACGGTGCCTCGGCGTGTG 900
DB |||||
QY 5987 CATGGCTCACTCGTAATTCCTATATACTAATACACCCCTATACGGTGCCTCGGCGTGTG 6046
DB |||||
QY 901 GACTTTGCCCTTGAGCTTGAGTTTTCGCAACCTTTACCCCGGTAAACCAATACGCGGTC 960
DB |||||
QY 6047 GACTTTGCCCTTGAGCTTGAGTTTTCGCAACCTTTACCCCGGTAAACCAATACGCGGTC 6106
DB |||||
QY 961 TCCCGTTATTCAGACACTGCTCGCACCGCTTCGTGCGGTGCGGACGGGACTGCCGAG 1020
DB |||||
QY 6107 TCCCGTTATTCAGACACTGCTCGCACCGCTTCGTGCGGTGCGGACGGGACTGCCGAG 6166
DB |||||
QY 1021 CTCACCAACACCGCTGCTACTCCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGT 1080
DB |||||
QY 6167 CTCACCAACACCGCTGCTACTCCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGT 6226
DB |||||
QY 1081 GTCGTTGAGATCGGCGCGGATAGCCCTTCAACCTTTCAACCTTCTGACACTCTGCTT 1140
DB |||||
QY 6227 GTCGTTGAGATCGGCGCGGATAGCCCTTCAACCTTTTAACTTCTGACACTCTGCTT 6286
DB |||||
QY 1141 GCGGCGCTGCCACAGAAATGATTTGTCGGCTGGTGGCAGCTGTCTACTCCCGTCC 1200
DB |||||
QY 6287 GCGGCGCTGCCACAGAAATGATTTGTCGGCTGGTGGCAGCTGTCTACTCCCGTCC 6346
DB |||||
QY 1201 GTTGTCTCAGCAATGGCGAGCGACTGTTTAAAGTTGTATACATCTGTAGAGATGCTCAG 1260
DB |||||
QY 6347 GTTGTCTCAGCAATGGCGAGCGACTGTTTAAAGTTGTATACATCTGTAGAGATGCTCAG 6406
DB |||||
QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTCGCTCGGAGATCTCGTGTGTTATT 1320
DB |||||
QY 6407 CAGGATAAGGGTATTGCAATCCCGCATGACATTCGCTCGGAGATCTCGTGTGTTATT 6466
DB |||||
QY 1321 CAGGATTATGATAACCAATGAAACAGATCGCGGACGCTTCTCCAGCCCATGCGC 1380
DB |||||
QY 6467 CAGGATTATGATAACCAATGAAACAGATCGCGGACGCTTCTCCAGCCCATGCGC 6526
DB |||||
QY 1381 CTTTCTCTGCTCTTCGAGCTAATGATGCTTTTGGCTCTCTCTACCGCTGCCAGTAT 1440
DB |||||
QY 6527 CTTTCTCTGCTCTTCGAGCTAATGATGCTTTTGGCTCTCTCTACCGCTGCCAGTAT 6586
DB |||||
QY 1441 GACCACTCACTTATGGCTCTTCGACTGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
DB |||||
QY 6587 GACCACTCACTTATGGCTCTTCGACTGCCAGTTTATGTTTCTGACTCTGTGACCTTG 6646
DB |||||
QY 1501 GTTATGTTGCAACCGCGCGAGCGCTGTCGCGTGCCTCGATTTGACCAAGGTACA 1560
DB |||||
QY 6647 GTTATGTTGCAACCGCGCGAGCGCTGTCGCGTGCCTCGATTTGACCAAGGTACA 6706
DB |||||
QY 1561 CTTGACGGTTCGCGCTCTCCACCATCCAGAGTACTCGAAGACCTTCTTCTGCTGCGG 1620
DB |||||
QY 6707 CTTGACGGTTCGCGCTCTCCACCATCCAGAGTACTCGAAGACCTTCTTCTGCTGCGG 6766
DB |||||
QY 1621 CTCGCGGTAAAGCTCTCTTTCTGGAGGAGGAGCAACAACTAAAGCCGGGTACCTTATAAT 1680
DB |||||
QY 6767 CTCGCGGTAAAGCTCTCTTTCTGGAGGAGGAGCAACAACTAAAGCCGGGTACCTTATAAT 6826
DB |||||
QY 1681 TATAACCACTGCTAGGACCAACTGCTTTGCGAGATGCCGCGGGACCGGTGCTG 1740
DB |||||
QY 6827 TATAACCACTGCTAGGACCAACTGCTTTGCGAGATGCCGCGGGACCGGTGCTG 6886
DB |||||
QY 1741 ATTTCACCTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB |||||
QY 6887 ATTTCACCTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6946
DB |||||
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGGACTACCTTCCCGCGCC 1860
DB |||||
QY 6947 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGGACTACCTTCCCGCGCC 7006
DB |||||
QY 1861 CATACTTTTATGATTTCTGCGCAGAGTGCAGCGCTTGGCTTGAAGGCTGCGCTTTC 1920
DB |||||
QY 7007 CATACTTTTATGATTTCTGCGCAGAGTGCAGCGCTTGGCTTGAAGGCTGCGCTTTC 7066
DB |||||

QY 1921 CAGTCTACTGTGCTGAGCTTCAGCGCTTAAAGATGAAGTGGTAAAACTCGGAGTTG 1980
DB |||||
QY 7067 CAGTCTACTGTGCTGAGCTTCAGCGCTTAAAGATGAAGTGGTAAAACTCGGAGTTA 7126
DB |||||
QY 1981 TAG 1983
DB 7127 TAG 7129
|||
RESULT 14
AAQ86592
ID AAQ86592 standard; DNA; 1984 BP.
XX
AC AAQ86592;
XX
AC
XX 25-MAR-2003 (revised)
DT 28-SEP-1995 (first entry)
XX
XX Hepatitis E virus ORF2.
DE
XX HEV; ORF2; antigen; vaccine; immunogen; ss.
XX
XX Hepatitis E virus.
FH Key Location/Qualifiers
FT CDS 1..1983
FT /*tag= a
XX
PN W09508632-A1.
XX
PD 30-MAR-1995.
XX
PF 23-SEP-1994; 94WO-AU0000572.
XX
PR 24-SEP-1993; 93AU-00001423.
PR 15-DEC-1993; 93AU-00002964.
XX
XX (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
XX Anderson DA, Locarnini SA, Torresi J, Li F, Hui Z;
PI
XX WPI; 1995-139601/18.
DR P-PSDB; AAR70323.
XX
PT Antigens of hepatitis E virus (HEV) - selectively immuno:reactive to
PT convalescent and/or acute phase circulating antibodies to HEV.
XX
PS Disclosure; Page 39-41; 78pp; English.
XX
CC RNA from an HEV strain isolated in the Xinjiang region of China was
CC subjected to RT-PCR to obtain fragments corresp. to ORF2 (given in
CC AAQ86592) and ORF3 (AAQ86593) that encoded antigenic proteins (AAR70323-
CC 24, respectively). DNA fragments were manipulated into pGEX vectors for
CC production of GST fusion proteins in E. coli. Applications include HEV
CC infection diagnosis, therapy and vaccine development. (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
SQ Sequence 1984 BP; 347 A; 666 C; 481 G; 490 T; 0 U; 0 Other;
Query Match 91.0%; Score 1805.4; DB 2; Length 1984;
Best Local Similarity 94.4%; Pred No. 0;
Matches 1872; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTTTGCCTATGCTGCCCGGCA 60
DB 1 ATGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTTTGCCTATGCTGCCCGGCA 60
QY 61 CGCCCGGTGAGCGTCTGCGCGCGCGCTGCGGCGCGCAGCGGCTTCCGCGGTG 120
DB 61 CGCCCGGTGAGCGTCTGCGCGCGCGCTGCGGCGCGCAGCGGCTTCCGCGGTG 120
QY 121 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTTCACCAAC 180
DB |||||

Db 121 TTCTGGGTGACGGGTTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Qy 181 CCTTCGCCCCCGATGTACCGCTGCGCGGGGGTGGACCTCGTGTTCGCCAACCGGC 240
Db 181 CCTTCGCCCCCGATGTACCGCTGCGCGGGGGTGGACCTCGTGTTCGCCAACCGGC 240
Qy 241 CGACCACTCGGCTCGGCTGGGCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 241 CGACCACTCGGCTCGGCTGGGCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Qy 301 AGACCTACCACTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Db 301 AGACCTACCACTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Qy 361 CGAGTGCCTGATGTGACCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Db 361 CGAGTGCCTGATGTGACCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Qy 421 TCTCCCTTACCTCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
Db 421 TCTCCCTTACCTCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
Qy 481 AGTCCGCTTTACCCCTTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Db 481 AGCCCACTTCTACCCCTTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Qy 541 AATTATGCCAGTACCGGGTTGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Db 541 AATTATGCCAGTACCGGGTTGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Qy 601 GCTGTCCGCGGTTACGCGCATCTCCATCTCTTCTGGCCACAGACCAACCCCGACG 660
Db 601 GCTGTCCGCGGTTACGCGCATCTCCATCTCTTCTGGCCACAGACCAACCCCGACG 660
Qy 661 TCCGCTTGATATGAATCAATAACCTCGACGAGTGTTCGTATTTAGTCCAGCCGCGGCA 720
Db 661 TCCGCTTGATATGAATCAATAACCTCGACGAGTGTTCGTATTTAGTCCAGCCGCGGCA 720
Qy 721 GCCTTGAGCTTGTGATCCCAAGTGGAGGCTACCTCTGCTGTGTATGTATGTTCGATA 780
Db 721 GCCTTGAGCTTGTGATCCCAAGTGGAGGCTACCTCTGCTGTGTATGTATGTTCGATA 780
Qy 781 GTCCGAGACCTTGCGGGTGTGAGGAGGCTACCTCTGCTGTGTATGTATGTTCGATA 840
Db 781 GTCCGAGACCTTGCGGGTGTGAGGAGGCTACCTCTGCTGTGTATGTATGTTCGATA 840
Qy 841 CATGGCTCACTCGTAAATTCCTATATAACCTTACCTATACCGGTGCCCTCGGCTGTG 900
Db 841 CATGGCTCACTCGTAAATTCCTATATAACCTTACCTATACCGGTGCCCTCGGCTGTG 900
Qy 901 GACTTTGCCCTTGAGCTTGATTTGCGAACCTTACCCCGGTACACCAATACCGGGTC 960
Db 901 GACTTTGCCCTTGAGCTTGATTTGCGAACCTTACCCCGGTACACCAATACCGGGTC 960
Qy 961 TCCGCTTATTCAGCACTGCTCGCCACCGCTTCTGCGGTGGGAGCGGAGCTGCCGAG 1020
Db 961 TCCGCTTATTCAGCACTGCTCGCCACCGCTTCTGCGGTGGGAGCTGCCGAG 1020
Qy 1021 CTCACCAACCGGCTGCTACCCGCTTATGAAGACCTCTATTTATAGTACTAATGGT 1080
Db 1021 CTTACCAACCGGCTGCTACCCGCTTATGAAGACCTCTATTTATAGTACTAATGGT 1080
Qy 1081 GTCCGTGAGATCGCGCGGATAGCCTCACCTGTTCACCTTGCATGCTGACCTCTGCT 1140
Db 1081 GTCCGTGAGATCGCGCGGATAGCCTCACCTGTTCACCTTGCATGCTGACCTCTGCT 1140
Qy 1141 GCGCGCTTCGCCACAGAAATTCGTCGCGTGGTGGCCAGCTGTTCTTACTCCCGTCCC 1200
Db 1141 GCGCGCTTCGCCACAGAAATTCGTCGCGTGGTGGCCAGCTGTTCTTACTCTCGTCCC 1200
Qy 1201 GTTGTCTCAGCAATGGCGAGCGGCTGTTAAGTTGTATATCATCTGTAGAGAAATGCTCAG 1260
Db 1201 GTTGTCTCAGCAATGGCGAGCGGCTGTTAAGTTGTATATCATCTGTAGAGAAATGCTCAG 1260

Qy 1261 CAGGATAAGGGTATTGCAATCCGCGATGACATTGACCTCGGAGAAATCTCGTGTGGTTATT 1320
Db 1261 CAGGATAAGGGTATTGCAATCCGCGATGACATTGACCTCGGAGAGTCTCGTGTAGTTATT 1320
Qy 1321 CAGGATTATGATAACCAACATGAACAAAGATCGCGCCGAGCCGCTTCTCAGGCCCATCGCGC 1380
Db 1321 CAGGATTATGACAAACCAACATGAGCAGGACCGACCGACACCTTCCCGCAGCCCATCGCGC 1380
Qy 1381 CCTTTCTGTCTGCTTCGAGCTAATGATGTGCTTTGCTCTCTCAGCGCTGCGGAGTAT 1440
Db 1381 CCTTTTCTGTCTGCTTCGAGCTAATGATGTGCTTTGCTCTCTCAGCGCTGCGGAGTAT 1440
Qy 1441 GACCACTCCACTTTATGGCTCTTTCGACTGGCCCAAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 1441 GACCACTCCACTTTATGGCTCTTTCGACTGGCCCAAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Qy 1501 GTTAAATGTTGCGACCGCGCGAGCGGCTTTCGCGTTCGATTCGATTCGACCAAGGTCACA 1560
Db 1501 GTTAAATGTTGCGACCGCGCGAGCGGCTTTCGCGTTCGATTCGACCAAGGTCACA 1560
Qy 1561 CTTGACGCTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCTGCGG 1620
Db 1561 CTTGATGGTGGCCCTTTCACCATCCAGCAGTATTCAAAGACCTTCTTTGCTCTGCGG 1620
Qy 1621 CTCGCGGTAAAGCTCTCTTTCTGGAGGCGAGGACCAACTAAAGCCGCGGTACCCCTTATAAT 1680
Db 1621 CTCGCGGTAAAGCTCTCTTTCTGGAGGCGAGGTACTACTAAAGCCGCGGTACCCCTTATAAT 1680
Qy 1681 TATAACCACTCTGCTAGCAGCAACTGCTGTTGAGAAATCGCGCTGCGCATCGGGTTGCT 1740
Db 1681 TATAACCACTCTGCTAGCAGCAACTGCTGTTGAGAAATCGCGCTGCGCATCGGGTTGCT 1740
Qy 1741 ATTTCCACTTTACCACTAGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1800
Db 1741 ATTTCCACTTTACCACTAGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1800
Qy 1801 TTAGCCCCCCTCTGCGCTAGCAATGCTTTGAGGATACCTTTGAGTACCTCCCTGCCCGGCC 1860
Db 1801 TTAGCCCCCCTCTGCGCTAGCAATGCTTTGAGGATACCTTTGAGTACCTCCCTGCCCGGCC 1860
Qy 1861 CATACTTTGATGATTTCTGCCAGAGTGGCGCCCTTGGCCCTTCAAGGCTGCGCTTTC 1920
Db 1861 CATACTTTGATGATTTCTGCCAGAGTGGCGCCCTTGGCCCTTCAAGGCTGCGCTTTC 1920
Qy 1921 CAGTCTACTCTCGCTGAGCTTCAGCGCTTAAGATGAAGTGGTAAAGCTCGGAGTTG 1980
Db 1921 CAGTCTACTCTCGCTGAGCTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCG 1980
Qy 1981 TAG 1983
Db 1981 TAG 1983

RESULT 15

ABL60053
ID ABL60053 standard; DNA; 1990 BP.

AC ABL60053;

XX
XX 16-AUG-2002 (first entry)

XX
XX Hepatitis E virus ORF-2 DNA #1.

XX
XX Hepatitis E virus; ORF-2; HEV; virucide; immunisation; gene; ds.

XX
XX Hepatitis E virus.

XX
XX WO200240681-A1.

XX
XX 23-MAY-2002.

XX
XX 30-SEP-2001; 2001MO-CN001469.

Db 1741 ATTTCACATTACACCACTAGCCTGGGTGCTGGTCCGCTCTCTATTTCGGGGTTGCTGTT 1800
QY 1801 TTAGCCCCCACACTCTGCGCTAGCATTTGCTTGAGGATACCTTGGACTACCTGCCCCGGCC 1860
Db 1801 TTAGCCCCCACACTCCGCGCTAGCATTTGCTTGAGGATACCATGGACTACCTGCCCCGGCC 1860
QY 1861 CATACTTTTGATGATTTCTGCCAGAGTGCCGCCCTTGCGCTTCAGGGCTGGCTTTC 1920
Db 1861 CATACTTTCGATGACTTCTGCCGGAGTGCCGCCCTTGCGCTTCAGGGCTGGCTTTT 1920
QY 1921 CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGTAAACTCGGGAGTTG 1980
Db 1921 CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGTAAACTCGGGAGTTA 1980
QY 1981 TAG 1983
Db 1981 TAG 1983

Search completed: August 21, 2004, 11:00:11
Job time : 775.579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 10:14:02 ; Search time 5116.87 seconds
(without alignments)
11572.836 Million cell updates/sec

Title: US-09-851-410a-6_COPY_5147_7129

Perfect score: 1983

Sequence: 1 ATGCGCCCTCGGCTATTTT.....GTAAACTCGGAGTTGTAG 1983

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gesl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67.2	3.4	925	29	CNS0091P
C 2	67.2	3.4	1201	9	AL578906
C 3	63.6	3.2	1201	13	AL578906
C 4	62.2	3.1	1103	13	AL578906

C 5	60.2	3.0	1201	13	BX360624
C 6	60.2	3.0	1839	29	CG747711
C 7	60	3.0	1057	13	BX349688
C 8	59.6	3.0	844	29	CNS0052P
C 9	59.6	3.0	1201	13	BX446579
C 10	59.2	3.0	839	29	CNS004NB
C 11	59	3.0	1793	29	CG754612
C 12	58.4	2.9	776	29	CNS010RY
C 13	58	2.9	932	29	CNS0072Q
C 14	58	2.9	1174	14	CD508311
C 15	58	2.9	1213	12	BM560246
C 16	57.6	2.9	1695	28	CC290874
C 17	57.4	2.9	1135	28	BZ574895
C 18	57.4	2.9	1970	29	CG748837
C 19	57.2	2.9	807	13	BX464554
C 20	57.2	2.9	935	29	CNS006XK
C 21	57	2.9	1136	14	CD507542
C 22	56.8	2.9	717	29	AG135089
C 23	56.4	2.8	946	29	AG080623
C 24	56.4	2.8	1000	13	BX407619
C 25	56.4	2.8	1061	13	BX391246
C 26	55.8	2.8	1007	28	BZ569268
C 27	55.8	2.8	1157	28	CC219063
C 28	55.6	2.8	964	29	CNS003WG
C 29	55.6	2.8	1016	13	BQ681552
C 30	55.6	2.8	1094	14	CD509997
C 31	55.6	2.8	1193	13	BX421125
C 32	55.4	2.8	1137	29	AG078502
C 33	55.4	2.8	1686	28	CC189908
C 34	55.2	2.8	995	14	CA788818
C 35	55.2	2.8	1032	13	BQ680891
C 36	55.2	2.8	1356	29	CG754548
C 37	54.8	2.8	1362	12	BM556092
C 38	54.6	2.8	948	13	BU953470
C 39	54.6	2.8	1107	28	BZ051753
C 40	54.6	2.8	1167	13	BQ061197
C 41	54.4	2.7	703	29	AG167751
C 42	54.4	2.7	739	29	AG080613
C 43	54.4	2.7	982	13	BX415111
C 44	54.4	2.7	983	13	BU838409
C 45	54.4	2.7	1036	9	AL550172

ALIGNMENTS

RESULT 1	CNS0091P	925 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
DEFINITION	BAC19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL053013				
VERSION	AL053013.1				
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Ephyridioidea; Drosophilidae; Drosophila.				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseogawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of				

/mol_type="mRNA"		vector."	
/db_xref="taxon:9606"			
/clone="CS0D1072YL05"			
/issue_type="PLACENTA COT 25-NORMALIZED"			
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"			
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."			
ORIGIN			
Query Match 3.0%; Score 60.2; DB 13; Length 1201;			
Best Local Similarity 9.2%; Pred. No. 0.0052;			
Matches 25; Conservative 163; Mismatches 85; Indels 0; Gaps 0;			
QY	65	CCGGTCAGCGCTGCGCGCGCTGCTGGCGCGCAGCGCGGTTCGGCGGTGTTCT	124
DB	1178	SBSTTTTTSBTSBSSSSSBSTTTTSSSSSBSTTSBSSSSSBSTTSBTTSTBS	1119
QY	125	GGGGTGACCGGTTGATCTCAGCCCTTCGCAATCCCTATATTCACCAACCAACCCCT	184
DB	1118	TTTTKSVTTCSTBTMSKTTSTCSYTTCTVTSYCSBECSTTSCCTTCTCTCSV	1059
QY	185	TGCCCCCGATGCACGCTCGCGCGCGGCTGACCTCGTGTTCGCCAACCCCGCGAC	244
DB	1058	STSSBECBSBTTTSSSSSSSSSSTTSBBSBSSSSSSSBSSSSSSSSSSSSSSS	999
QY	245	CACGCGCTCGCTGCGGTGACACAGCCAGCGCCCGCGCTTGCCTCAGCTCGTAGAC	304
DB	998	SSSSBBSBTSBSSSSSBSTSSSVSSSSSSSSSSSSSBSCSBBSBTSBSSSSSSS	939
QY	305	CTACCACAGCTGGGCGCGCGCTAACCGCG 337	
DB	938	SSSSSASSSSSSSSSSTSKWSSTASSSSSV 906	
RESULT 6			
CG747711/c			
LOCUS			
DEFINITION			
P041-2-E02.2a Ppa EcoRI BAC Library Pristionchus pacificus genomic,			
genomic survey sequence.			
CG747711			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Pristionchus pacificus			
Pristionchus pacificus			
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;			
Neodiplogasteridae; Pristionchus.			
1 (bases 1 to 1839)			
Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,			
Buntjer,J., van der Meulen,M. and Sommer,R.J.			
An integrated physical and genetic map of the nematode Pristionchus			
pacificus			
Mol. Genet. Genomics 269 (5), 715-722 (2003)			
JOURNAL			
MEDLINE			
PUBMED			
COMMENT			
Contact: Sommer RJ			
Evolutionary Biology			
Max-Planck-Institute for Developmental Biology			
Spemannstr. 37-39, Tuebingen D-72076, Germany			
Tel: 00497071601371			
Fax: 00497071601498			
Email: ralf.sommer@tuebingen.mpg.de			
Class: BAC ends.			
Location/Qualifiers			
1..1839			
/organism="Pristionchus pacificus"			
/mol_type="genomic DNA"			
/strain="California"			
/db_xref="taxon:54126"			
/clone_lib="Ppa EcoRI BAC Library"			
/note="The library was generated by a partial digest of			
the genomic DNA with EcoRI and cloning into the BAC			
FEATURES			
source			

```

source
1. .1057
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI063YK05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match      3.0%; Score 60; DB 13; Length 1057;
Best local Similarity 32.1%; Pred. No. 0.0055;
Matches 122; Conservative 77; Mismatches 181; Indels 0; Gaps 0;

QY 150 CTTGCGAATCCCTATATTCATCAACAACCCCTTCGCCGCCCATGTACCGTGGCGC 209
Db 692 CATCGCTATTAAACATGGTATGCGGYCTTSCGCGGCCCCCCCCCGCGCGCGCGCGC 633

QY 210 CGGCGCTGGACCTCGTGTTCGCCAACCGCCGCGACCACTCGGCTCGCTTGGCGCTGACCA 269
Db 632 GGCCSCSCSCSGSGSCSCSGSGSCGCGGCCCTCCSCSCSGSGSGCGCCSCCC 573

QY 270 GGCCGAGCGCCCGCGCGCTTGCCTCACTGCTGAGACCTACCACTGCGCGCGCGCGCT 329
Db 572 CCCCCSCGCCSCGSSSCSCSCSSSSSCGSCSSCSGCGGCCCCSCSSSCSSSS 513

QY 330 AACCGGCGTCTCGGCGCCATGACACCCGCGCAGTGCCTGATGTCACATCCCGCGCGC 389
Db 512 SCGSSSSSGSSCGCGSGCGSCSCCCCCSCCGSCSCSCSCSCSCSCSCSCSCSCSC 453

QY 390 CATCTTGGCGCGCAGTATACCTATCAACATCTCCCTTACTCTTCTCGTGGCACCGG 449
Db 452 CGSCCGSCSCSCCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 393

QY 450 CACTAACCTGGTCTTTATGCGCGCCCTCTTAGTCGCGCTTTTACCCCTTCAGGACGCGAC 509
Db 392 CGCCCGCCSCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 333

QY 510 CAATACCCATATAATGGCCA 529
Db 332 MSCCTMMSMCCATNAGGCTA 313

CNS0052P      844 bp      DNA      linear      GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACRilP16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL056652
AL056652.1   GI:4932342
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 844)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of

```

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPcI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
 source 1..844
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR11P16"
 /clone_lib="RPcI-98"
 /note="end : TET3"

ORIGIN

Query Match 3.0%; Score 59.6; DB 29; Length 844;
 Best Local Similarity 19.9%; Pred. No. 0.0061;
 Matches 59; Conservative 127; Mismatches 111; Indels 0; Gaps 0;

QY 21 GTTCTGCTCCTCATGTTTGGCTATGTGCCCGCGCACCGCCGCTAGCGCTCGG 80
 Db 678 BTTCTWHTNSNTYTMVTYTTTHTTNTCCYCCCCYCCCBYBCNTSCSYYSB 619
 QY 81 CCGCCGCTGCGCGCGCAGCGCGGTCCGCGCGTGTCTCTGGGTGACCGGTTGA 140
 Db 618 CCGCSSSSSSSSGSSBYSCCBCCCBGTCGYBVCBSBSSTBSVSSGSSSTSSGTG 559
 QY 141 TTCTACAGCCCTTGGCAATCCCTTATCATCAACCAACCCCTTCGCCCCGATGTCAC 200
 Db 558 CKCBSSSSBBSBSSCCYCSYCTCYSSCBYCCBSGCBTBSYBSCSSGBTGBSSB 499
 QY 201 CGTTCGCGCGGGCTGAGACTGTGTTGCCAAACCGCCCGACACTCGTCCGCTTG 260
 Db 498 SGGKSGKSGKGGKSGKSGKGYCYBTYCCSBSBSCSSGSSSTSKCSGSSSTKGS 439
 QY 261 GCGTGACCAAGCCAGCGCCCGCGTTCCTCACGTCGTACACTACACAGCTGG 317
 Db 438 SBBGGSSSKTSGSSCBSSGSSSTGSSGSKSYSTSSSSGCGTGKNTTTCGVS 382

EX446579 1201 bp mRNA linear EST 22-MAY-2003
 EX446579 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
 CL0BB002ZF10 3-PRIME, mRNA sequence.
 EX446579
 EX446579.1 GI:31033746
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1201)
 Li W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 7325.f For
 more information about this cluster, see
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CL0BB002ZF10FP1&cluster=7325.f)
 cgi-bin/cluster.cgi?seq=CL0BB002ZF10FP1&cluster=7325.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CL0BB002ZF10FP1.
 Location/Qualifiers
 1..1201

FEATURES
 source 1..1201

High quality sequence stop: 172.
Location/Qualifiers
1. .1174

```
1. .1174
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Salinas river, CA"
/db_xref="taxon:69293",
/clone="CDA90-105"
/sex="mixed male and female"
/tissue_type="heads and internal organs combined"
/dev_stage="adult"
```

Match 2.9%; Score 58; DB 14; Length 1174;
 Local Similarity 30.7%; Pseud. No. 0.017;
 163; Conservative 0; Mismatches 368; Indels 0; Gaps 0;

```

132  CGGGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCCCTTGGCC 191
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195  CCNNNNNNNNNNNNCCNCNCCCCCNCNNNNCCNCCCCCNCNNNNCCCC 936
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
192  CGATGTACCGGTGCGGGGGCTGGACCTGTTGTGGCAAACCGCGCGGACACTCGG 251
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
935  CNCCNCCNNNNNCCCCCNCNNNNCNMCCCCCNCNNCCCCCGCCGCCCCNNNN 876
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
252  CTCGGTGTGGGTGACAGAGCCAGGGCCCCCGGTGTGCTCAGTCTAGACTTACCAC 311
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
875  CCGCCNCCNNNNNNNNNNCCCNCCNNNNCCNNNNCCCCCCCCCCCCCNCNNCCCC 816
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
112  AGCTCCGCGCGCGCGCGCTAAACCGCGGTGCTCCCGCCCATGACACCGCGCAGTGCCTGA 371
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

[illegible]

Blank Sheet

APPLICANT: Tam, Albert W.
 APPLICANT: Krawczynski, Krzysztof Z.
 APPLICANT: Yarbough, Patrice D.
 TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
 NUMBER OF SEQUENCES: 60
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 350 Cambridge Avenue, Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US 08/259,148A
 FILING DATE: 13-JUN-1994
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 822,335
 FILING DATE: 17-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 505,888
 FILING DATE: 05-APR-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 420,921
 FILING DATE: 13-OCT-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 367,486
 FILING DATE: 16-JUN-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 336,672
 FILING DATE: 11-APR-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 208,997
 FILING DATE: 17-JUN-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Sholtz, Charles K.
 REGISTRATION NUMBER: 38,615
 REFERENCE/POCKET NUMBER: 4600-0093.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2049 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 7
 US-08-259-148A-3

Query Match 100.0%; Score 1983; DB 1; Length 2049;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCGCCCTCGGCCTATTTCGTCTGCTCTCTCATGTTTTTTCCTATGCTGCGCGGCCA	60
Db	1	ATGGCGCCCTCGGCCTATTTCGTCTGCTCTCTCATGTTTTTTCCTATGCTGCGCGGCCA	60
Qy	61	CCGCGCGGTACCGCGTCTGCGCGCGGTCTGCGCGCGGTCCGCGCGGTGGT	120
Db	61	CCGCGCGGTACCGCGTCTGCGCGCGGTCTGCGCGCGGTCCGCGCGGTGGT	120
Qy	121	TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC	180
Db	121	TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC	180

Qy	181	CCCTTCGCCCGCCGATGTGTACCCCTCGCGCGCGGGCTGGACCTCGTGTTCGCCAACCCGCC	240
Db	181	CCCTTCGCCCGCCGATGTGTACCCCTCGCGCGCGGGCTGGACCTCGTGTTCGCCAACCCGCC	240
Qy	241	CGACCACTCGGCTCGGCTTGGCGTGACCAAGGCCAGCGCCCGCGCTTCGCTCAGCAACCCCG	300
Db	241	CGACCACTCGGCTCGGCTTGGCGTGACCAAGGCCAGCGCCCGCGCTTCGCTCAGCAACCCCG	300
Qy	301	AGACCTACCAAGCTGGGGCCCGCGGTAAACCGGGTCTCGGCTCCGCTCAGCAACCCCG	360
Db	301	AGACCTACCAAGCTGGGGCCCGCGGTAAACCGGGTCTCGGCTCCGCTCAGCAACCCCG	360
Qy	361	CCAGTGCCCTGTGTGAGTCTCCCGCGCGGCATCTTGGCGCGCAGTATTAACCTATCAACA	420
Db	361	CCAGTGCCCTGTGTGAGTCTCCCGCGCGGCATCTTGGCGCGCAGTATTAACCTATCAACA	420
Qy	421	TCCTCCCTTACCTCTTCGCTGGCCACCGGCATTAACCTGGTGTCTTTATGCGCCCTCTT	480
Db	421	TCCTCCCTTACCTCTTCGCTGGCCACCGGCATTAACCTGGTGTCTTTATGCGCCCTCTT	480
Qy	481	AGTCCGCTTTTACCCCTTCAAGGACCGCACCAATACCCATATAATGCGCCACGGAAGCTTCT	540
Db	481	AGTCCGCTTTTACCCCTTCAAGGACCGCACCAATACCCATATAATGCGCCACGGAAGCTTCT	540
Qy	541	AATTATGCCAGTACCGGGTTCGCGTGCCAGTATCCGTTACCGCCCGCTGTCCTCCCAAT	600
Db	541	AATTATGCCAGTACCGGGTTCGCGTGCCAGTATCCGTTACCGCCCGCTGTCCTCCCAAT	600
Qy	601	GCTGTGGCGGTTCAGCCATCTCCATCTCATCTCTGGCCACAGACCACCAACCCCGAG	660
Db	601	GCTGTGGCGGTTCAGCCATCTCCATCTCATCTCTGGCCACAGACCACCAACCCCGAG	660
Qy	661	TCCGTTGATATGAATTCAATTAACCTCGACGATGTTCGTATTTAGTCAGCCCGGCATA	720
Db	661	TCCGTTGATATGAATTCAATTAACCTCGACGATGTTCGTATTTAGTCAGCCCGGCATA	720
Qy	721	GCCTGTAGCTTGTGATCCCAAGTGAGGCGCTACACTATCGTAACCAAGGTGGCGTCC	780
Db	721	GCCTGTAGCTTGTGATCCCAAGTGAGGCGCTACACTATCGTAACCAAGGTGGCGTCC	780
Qy	781	GTCGAGACTCTCGGGTGGCTGAGGAGGAGCTACCTCTGCTGTCTGTTATGCTTGCATA	840
Db	781	GTCGAGACTCTCGGGTGGCTGAGGAGGAGCTACCTCTGCTGTCTGTTATGCTTGCATA	840
Qy	841	CATGGCTCACTCGTAAATTCCTATACTAATAACCCCTATACCGTGCCTCGGCGTGTG	900
Db	841	CATGGCTCACTCGTAAATTCCTATACTAATAACCCCTATACCGTGCCTCGGCGTGTG	900
Qy	901	GACTTTCGCTTGTGAGCTTGTGCTTTCGCAACCTTACCCCGGTAAACCAATACCGGCTC	960
Db	901	GACTTTCGCTTGTGAGCTTGTGCTTTCGCAACCTTACCCCGGTAAACCAATACCGGCTC	960
Qy	961	TCCGTTTATTCAGCACTGCTCGCCACCGCTTCGTGCGGTGCGGAGGAGTCCGAG	1020
Db	961	TCCGTTTATTCAGCACTGCTCGCCACCGCTTCGTGCGGTGCGGAGGAGTCCGAG	1020
Qy	1021	CTACCAACCAAGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTAAATGGT	1080
Db	1021	CTACCAACCAAGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTAAATGGT	1080
Qy	1081	GTCGCTGAGATCGCGCGCGGATAGCCCTCACCTGTTTCAACCTTGTGACACTCTGCTT	1140
Db	1081	GTCGCTGAGATCGCGCGCGGATAGCCCTCACCTGTTTCAACCTTGTGACACTCTGCTT	1140
Qy	1141	GGCGGCTGCGGACAGAAATTGTTTCGTGCGGTGCGGCGGAGTCTACTCCCGTCC	1200
Db	1141	GGCGGCTGCGGACAGAAATTGTTTCGTGCGGTGCGGCGGAGTCTACTCCCGTCC	1200
Qy	1201	GTTGTCTCAGCCAAATGGCGGAGCGGCTGTTTAAAGTTGTATACATCTGTAGAGATGCTCAG	1260
Db	1201	GTTGTCTCAGCCAAATGGCGGAGCGGCTGTTTAAAGTTGTATACATCTGTAGAGATGCTCAG	1260

1261 CAGGATAAGGGTATTGCAATCCGCATGACATTGACCTCGGAGAACTCTGCTGGTTATT 1320
 1261 CAGGATAAGGGTATTGCAATCCGCATGACATTGACCTCGGAGAACTCTGCTGGTTATT 1320
 1321 CAGGATTATGATAACCAACATGAACAAAGATCGGCGAGCCCTTCTCCAGCCCCATCGGCG 1380
 1321 CAGGATTATGATAACCAACATGAACAAAGATCGGCGAGCCCTTCTCCAGCCCCATCGGCG 1380
 1381 CTTTCTCTGCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGGAGTAT 1440
 1381 CTTTCTCTGCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGGAGTAT 1440
 1441 GACCACTCACTTATGGCTCTTGCAGTCCGAGCCGAGTATGATGCTTCTGACCTGTCGCTTG 1500
 1441 GACCACTCACTTATGGCTCTTGCAGTCCGAGCCGAGTATGATGCTTCTGACCTGTCGCTTG 1500
 1501 GTTAATGTTGCGACCGGCGCGAGCGGCTTGGCCGCTCGATGACCAAGGTACCA 1560
 1501 GTTAATGTTGCGACCGGCGCGAGCGGCTTGGCCGCTCGATGACCAAGGTACCA 1560
 1561 CTTGACGGTGGCGCCCTCTCCACCATCCAGCACTACTCGAAGACCTTCTTTGCTCTGCGG 1620
 1561 CTTGACGGTGGCGCCCTCTCCACCATCCAGCACTACTCGAAGACCTTCTTTGCTCTGCGG 1620
 1621 CTCGCGGTAAAGCTCTCTTCTGGAGGCGAGGACCAACTAAAGCGGGTACCTTATAAT 1680
 1621 CTCGCGGTAAAGCTCTCTTCTGGAGGCGAGGACCAACTAAAGCGGGTACCTTATAAT 1680
 1681 TATAACCACTACTAGCGACCAACTGCTTGTGAGAAATGCCCGCGGCAACCGGTCGCT 1740
 1681 TATAACCACTACTAGCGACCAACTGCTTGTGAGAAATGCCCGCGGCAACCGGTCGCT 1740
 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTGCGGTT 1800
 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTGCGGTT 1800
 1801 TTAGCCCCCACTCTCGGTAGCATGCTTGGAGTACCTTGGACTACCTGCGCGCGCC 1860
 1801 TTAGCCCCCACTCTCGGTAGCATGCTTGGAGTACCTTGGACTACCTGCGCGCGCC 1860
 1861 CATACTTTTATGATTTCTCCAGAGTGGCGGCGGCTTGGGCTTCCAGGCTCGGCTTTC 1920
 1861 CATACTTTTATGATTTCTCCAGAGTGGCGGCGGCTTGGGCTTCCAGGCTCGGCTTTC 1920
 1921 CAGTCTACTGCTGAGCTTACGCGCTTAAAGTGAAGTGGTAAACTCGGAGTTG 1980
 1921 CAGTCTACTGCTGAGCTTACGCGCTTAAAGTGAAGTGGTAAACTCGGAGTTG 1980
 1981 TAG 1983
 1981 TAG 1983

RESULT 3
 US-08-484-054-3
 ; Sequence 3, Application US/08484054
 ; Patent No. 5770689
 ; GENERAL INFORMATION:
 ; APPLICANT: Reyes, Gregory R.
 ; APPLICANT: Bradley, Daniel W.
 ; APPLICANT: Twu, Ji-Shin
 ; APPLICANT: Purdy, Michael A.
 ; APPLICANT: Tam, Albert W.
 ; APPLICANT: Krawczynski, Krzysztof Z.
 ; TITLE OF INVENTION: Hepatitis B Virus Vaccine and Method
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dehlinger & Associates
 ; STREET: 350 Cambridge Avenue, Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,054
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 870,985
 FILING DATE: 20-APRIL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 822,335
 FILING DATE: 17-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 505,888
 FILING DATE: 05-APRIL-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 420,921
 FILING DATE: 13-OCTOBER-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 367,486
 FILING DATE: 16-JUNE-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 336,672
 FILING DATE: 11-APRIL-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 208,997
 FILING DATE: 17-JUNE-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Sholtz, Charles K.
 REGISTRATION NUMBER: 38,615
 REFERENCE/DOCKET NUMBER: 4600-0093.38
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2049 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 7
 US-08-484-054-3

Query Match 100.0%; Score 1983; DB 1; Length 2049;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGCTGCGCGCCA 60
 Db 1 ATGGCGCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGCTGCGCGCCA 60

QY 61 CCGCGCGCTCAGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 120
 Db 61 CCGCGCGCTCAGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 120

QY 121 TTCTGGGGTGACCGGGTGTGATTCTCAGCCCTTTCGAAATCCCTATATATCAACCAAC 180
 Db 121 TTCTGGGGTGACCGGGTGTGATTCTCAGCCCTTTCGAAATCCCTATATATCAACCAAC 180

QY 181 CCCTTCCCGCCCGATGTCAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCC 240
 Db 181 CCCTTCCCGCCCGATGTCAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCC 240

QY 241 CGACCACTCGGCTCGGCTGCGGTGACAGCGCGCGCGCGCGCGCGCGCGCGT 300
 Db 241 CGACCACTCGGCTCGGCTGCGGTGACAGCGCGCGCGCGCGCGCGCGCGT 300

301 AGACCTACCAAGCTGGGGCGCGCGCTAAACCGCGTGGCTCCGCGCCCATGACACCCCG 360
Db |||||
301 AGACCTACCAAGCTGGGGCGCGCGCTAAACCGCGTGGCTCCGCGCCCATGACACCCCG 360
Qy |||||
361 CCAGTGCCTGATGCTGACCTCCGCGCGCGCGCTTGGCGCGGAGATTAACCTATCAACA 420
Db |||||
361 CCAGTGCCTGATGCTGACCTCCGCGCGCGCGCTTGGCGCGGAGATTAACCTATCAACA 420
Qy |||||
421 TCTCCCTTTACCTCTTCCGCTGGCCACCGGCACTTAACCTGCTTCTTATGCCGCCCTCTT 480
Db |||||
421 TCTCCCTTTACCTCTTCCGCTGGCCACCGGCACTTAACCTGCTTCTTATGCCGCCCTCTT 480
Qy |||||
481 AGTCCGCTTTTACCTCTTCCGCTGGCCACCGGCACTTAACCTATTAATGCCCGCAAGCTTCT 540
Db |||||
481 AGTCCGCTTTTACCTCTTCCGCTGGCCACCGGCACTTAACCTATTAATGCCCGCAAGCTTCT 540
Qy |||||
541 AATTATGCCAGTACCGGGTGGCGGCTGCGCAATCCGTTACCGCGCTGGTCCCAAT 600
Db |||||
541 AATTATGCCAGTACCGGGTGGCGGCTGCGCAATCCGTTACCGCGCTGGTCCCAAT 600
Qy |||||
601 GCTGTCCGCGGTTACCGCACTCTCCATCTCAATCTGSCCAGACACCAACCCCGGCG 660
Db |||||
601 GCTGTCCGCGGTTACCGCACTCTCCATCTCAATCTGSCCAGACACCAACCCCGGCG 660
Qy |||||
661 TCCGTTGATATGAATTCATTAACCTCGACGATGTTGATTTAGTCCAGCCCGGCATA 720
Db |||||
661 TCCGTTGATATGAATTCATTAACCTCGACGATGTTGATTTAGTCCAGCCCGGCATA 720
Qy |||||
721 GCCTCTGAGCTGTGATCCCAAGTAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
Db |||||
721 GCCTCTGAGCTGTGATCCCAAGTAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
Qy |||||
781 GTCGAGACTCTGGGGTGGCTGAGGAGGAGGCTACTCTGTGCTGTTATGCTTTGCGATA 840
Db |||||
781 GTCGAGACTCTGGGGTGGCTGAGGAGGAGGCTACTCTGTGCTGTTATGCTTTGCGATA 840
Qy |||||
841 CATGGCTCACTCGTAAATTCCTATATAACACCCCTATACCGGTGCCCTCGGGCTGTTG 900
Db |||||
841 CATGGCTCACTCGTAAATTCCTATATAACACCCCTATACCGGTGCCCTCGGGCTGTTG 900
Qy |||||
901 GACTTTGGCTTGAGCTTGAGTTTGGCACTTACCCCGGTAACCAATACGCGGCTC 960
Db |||||
901 GACTTTGGCTTGAGCTTGAGTTTGGCACTTACCCCGGTAACCAATACGCGGCTC 960
Qy |||||
961 TCCGCTTATTCAGCACTGCTCGCCACCGCTTGTGCGGCTGGCGGAGGAGTCCCGAG 1020
Db |||||
961 TCCGCTTATTCAGCACTGCTCGCCACCGCTTGTGCGGCTGGCGGAGGAGTCCCGAG 1020
Qy |||||
1021 CTCACCAACCGGCTGTACCGCTTTATGAGGAGCTCTATTTTACTAGTACTAATGGT 1080
Db |||||
1021 CTCACCAACCGGCTGTACCGCTTTATGAGGAGCTCTATTTTACTAGTACTAATGGT 1080
Qy |||||
1081 GTCGGTGAGATCGCGCGGAGTACCTTCACTGTTTCAACCTTGTGAGCACTCTGTT 1140
Db |||||
1081 GTCGGTGAGATCGCGCGGAGTACCTTCACTGTTTCAACCTTGTGAGCACTCTGTT 1140
Qy |||||
1141 GCGGCTTCCGCGACAGAAATGATTTGCTGCGCTGGTGGCGGAGCTTCTACTCCCGTCC 1200
Db |||||
1141 GCGGCTTCCGCGACAGAAATGATTTGCTGCGCTGGTGGCGGAGCTTCTACTCCCGTCC 1200
Qy |||||
1201 GTTGTCTCAGCAATGCGGAGCGGCTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db |||||
1201 GTTGTCTCAGCAATGCGGAGCGGCTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Qy |||||
1261 CAGGATAGGTTATGCAATCCCGCATGACATGACCTCGGAGAAATCTCGTGTGTTATT 1320
Db |||||
1261 CAGGATAGGTTATGCAATCCCGCATGACATGACCTCGGAGAAATCTCGTGTGTTATT 1320
Qy |||||
1321 CAGGATATGATTAACCAACATGAACATCGCGGAGCGCTTCTCCAGCCCAATCGGC 1380
Db |||||
1321 CAGGATATGATTAACCAACATGAACATCGCGGAGCGCTTCTCCAGCCCAATCGGC 1380
Qy |||||
1381 CCTTTCTGTCTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCCGAGTAT 1440

1381 CTTTTCTGTCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGTGCAGATAT 1440
Qy |||||
1441 GACAGTCCACTATAGCTTCTCGACTGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db |||||
1441 GACAGTCCACTATAGCTTCTCGACTGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Qy |||||
1501 GTTAATGTTGACACCGGCGGCGGAGCGCTTGGCCGGTCCGCTCGATTGGACCAAGGTACA 1560
Db |||||
1501 GTTAATGTTGACACCGGCGGCGGAGCGCTTGGCCGGTCCGCTCGATTGGACCAAGGTACA 1560
Qy |||||
1561 CTTGACGCTGCGCCCTCTCTCCACCACTCCAGCAGTACTCCAGAGACCTTCTTTGTCCTGCG 1620
Db |||||
1561 CTTGACGCTGCGCCCTCTCTCCACCACTCCAGCAGTACTCCAGAGACCTTCTTTGTCCTGCG 1620
Qy |||||
1621 CTTCCGCGGTAAGCTCTCTTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCTTTATAAT 1680
Db |||||
1621 CTTCCGCGGTAAGCTCTCTTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCTTTATAAT 1680
Qy |||||
1681 TATAACACCACTGTAGCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db |||||
1681 TATAACACCACTGTAGCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Qy |||||
1741 ATTTCACACTTACACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db |||||
1741 ATTTCACACTTACACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Qy |||||
1801 TTAGCCCCCACTCTGCGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db |||||
1801 TTAGCCCCCACTCTGCGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Qy |||||
1861 CATACTTTTGTGATGATTTCTGCGCAGAGTCCGCCCTTAAAGATGAAGTGGTAAAGTTCGGAGTTG 1920
Db |||||
1861 CATACTTTTGTGATGATTTCTGCGCAGAGTCCGCCCTTAAAGATGAAGTGGTAAAGTTCGGAGTTG 1920
Qy |||||
1921 CAGTCTACTGCTGCTGAGCTTACGCGCTTACGCGCTTAAAGATGAAGTGGTAAAGTTCGGAGTTG 1980
Db |||||
1921 CAGTCTACTGCTGCTGAGCTTACGCGCTTACGCGCTTAAAGATGAAGTGGTAAAGTTCGGAGTTG 1980
Qy |||||
1981 TAG 1983
Db |||||
1981 TAG 1983

RESULT 4

US-07-876-941A-3
; Sequence 3, Application US/07876941A
; Patent No. 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; TITLE OF INVENTION: Antibodies
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/876,941A
; FILING DATE: 01-MAY-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.33
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2049 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 7
US-07-876-941A-3

Query Match 100.0%; Score 1983; DB 2; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCTATTTGTTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
DB 1 ATGCGCCCTCGGCTATTTGTTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
QY 61 CCGCGCGGTGAGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCG 120
DB 61 CCGCGCGGTGAGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCG 120
QY 121 TTCTGGGGTGACCGGGTGTGTTCTCAGCCCTTCGCAATCCCTATATTCACCAAC 180
DB 121 TTCTGGGGTGACCGGGTGTGTTCTCAGCCCTTCGCAATCCCTATATTCACCAAC 180
QY 181 CCCTTCGCGCCCGGATGTACCGGTGCGCGCGGGGCTGGACCTCGTGTTCGCCAACCGCC 240
DB 181 CCCTTCGCGCCCGGATGTACCGGTGCGCGCGGGGCTGGACCTCGTGTTCGCCAACCGCC 240
QY 241 CGACCACTCGGCTCGGCTTGGCGTGACAGGCCCGCGCGCGCGGTGCGCTACGTCGT 300
DB 241 CGACCACTCGGCTCGGCTTGGCGTGACAGGCCCGCGCGCGGTGCGCTACGTCGT 300
QY 301 AGACCTACCACTGCGGCGCGCGCTAACCGCGGTGCGTCCGCGCCCATGACACCCCG 360
DB 301 AGACCTACCACTGCGGCGCGCGCTAACCGCGGTGCGTCCGCGCCCATGACACCCCG 360
QY 361 CCAGTGCGTATGTGATCTCCGCGGCGCCATCTTGGCGCGGAGTATTAACCTATCAACA 420
DB 361 CCAGTGCGTATGTGATCTCCGCGGCGCCATCTTGGCGCGGAGTATTAACCTATCAACA 420
QY 421 TCTCCCTTACCTCTCGGTGGCCACCGGCACTAACCTGTTCTTTATGCGCGCCCTCT 480
DB 421 TCTCCCTTACCTCTCGGTGGCCACCGGCACTAACCTGTTCTTTATGCGCGCCCTCT 480
QY 481 AGTCGGCTTTTACCCCTTCAGGAGCGGCAACCAATACCCCATATAATGGCCACGGAGCTCT 540

DB 481 AGTCGGCTTTTACCCCTTCAGGAGCGGCAACCAATACCCATATAATGGCCACGGAGCTCT 540
QY 541 AATTATGCCAGTACCGGGTGCCTGCGCAAAATCGGTACCGCGCGTGGTCCCAAT 600
DB 541 AATTATGCCAGTACCGGGTGCCTGCGCAAAATCGGTACCGCGCGTGGTCCCAAT 600
QY 601 GCTGTCGGCGGTACGCCATCTCCATCTCATCTGGCCACAGACACCAACCCCGGAGC 660
DB 601 GCTGTCGGCGGTACGCCATCTCCATCTCATCTGGCCACAGACACCAACCCCGGAGC 660
QY 661 TCCGTTGATATGAATTTCAATAACCTCGACGAGTGTGTTATTTTAGTCCAGCCCGGCA 720
DB 661 TCCGTTGATATGAATTTCAATAACCTCGACGAGTGTGTTATTTTAGTCCAGCCCGGCA 720
QY 721 GCTCTGAGCTTGTGATCCAGTGCCTACATATCGTAAACCAAGCTGGCGCTCC 780
DB 721 GCTCTGAGCTTGTGATCCAGTGCCTACATATCGTAAACCAAGCTGGCGCTCC 780
QY 781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGTACCTCTGGTCTTGTATGCTTTGCATA 840
DB 781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGTACCTCTGGTCTTGTATGCTTTGCATA 840
QY 841 CATGGCTCATCTCGTAAATTCCTATATACTAATAACCCCTATACCGGTGCCCTCGGGCTG 900
DB 841 CATGGCTCATCTCGTAAATTCCTATATACTAATAACCCCTATACCGGTGCCCTCGGGCTG 900
QY 901 GACTTTGCCCTTGTGCTTTCGCAACCTTTACCCCGGTAAACCAATACCGCGGTC 960
DB 901 GACTTTGCCCTTGTGCTTTCGCAACCTTTACCCCGGTAAACCAATACCGCGGTC 960
QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTGCGGTGCGGACGCGAGCTGCGAG 1020
DB 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTGCGGTGCGGACGCGAGCTGCGAG 1020
QY 1021 CTCACCAACCGCTGCTACCGCTTTATGAAGGAGCTCTATTTTACTAGTACTAATGGT 1080
DB 1021 CTCACCAACCGCTGCTACCGCTTTATGAAGGAGCTCTATTTTACTAGTACTAATGGT 1080
QY 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
DB 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
QY 1141 GCGCGCTGCGGACAGAAATGATTTGCTGCGGTGCGGCGAGCTGTTCTACTCCGTC 1200
DB 1141 GCGCGCTGCGGACAGAAATGATTTGCTGCGGTGCGGCGAGCTGTTCTACTCCGTC 1200
QY 1201 GTTGTCTCAGCCAAATGCGGAGCGGACTGTTAAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
DB 1201 GTTGTCTCAGCCAAATGCGGAGCGGACTGTTAAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
QY 1261 CAGGATAAGGGTATGCAATCCCGCATGACATTCGAGGAGTCTCGTGTGTTATT 1320
DB 1261 CAGGATAAGGGTATGCAATCCCGCATGACATTCGAGGAGTCTCGTGTGTTATT 1320
QY 1321 CAGGATTATGATAACCAACATGAACAGATCGGCGGAGCTTCTCAGCCCGCATCGCG 1380
DB 1321 CAGGATTATGATAACCAACATGAACAGATCGGCGGAGCTTCTCAGCCCGCATCGCG 1380
QY 1381 CTTTCTCTGTCTTCAGCTAATGATGCTTGGCTCTCTCTCAGCGTCCGAGTAT 1440
DB 1381 CTTTCTCTGTCTTCAGCTAATGATGCTTGGCTCTCTCTCAGCGTCCGAGTAT 1440
QY 1441 GACCACTGCTATGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
DB 1441 GACCACTGCTATGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
QY 1501 GTTAATGTCGACCGCGCGAGCGGTGCGCGGTGCGGTGCGTGTGATGTGAACCAAGTCA 1560
DB 1501 GTTAATGTCGACCGCGCGAGCGGTGCGCGGTGCGGTGCGTGTGATGTGAACCAAGTCA 1560
QY 1561 CTTGACCGGTGCGCGCTCTCCACCTACGAGTACTCGAGACCTTCTTGTCTGCGCG 1620

Db 1561 CTTGACGGTGGCCCCCTCTCCACCATCCAGAGTACTCGAAGACCTCTTTTGTCTGGCG 1620
QY 1621 CTCGGGGTAAGCTCTCTTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCTTTATAAT 1680
Db 1621 CTCGGGGTAAGCTCTCTTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCTTTATAAT 1680
QY 1681 TATAACACACTGTAGCGACCAACTGCTTGTGAGAAATCCGCGGACCGGGTCTGCT 1740
Db 1681 TATAACACACTGTAGCGACCAACTGCTTGTGAGAAATCCGCGGACCGGGTCTGCT 1740
QY 1741 ATTTCACACTACACACTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1800
Db 1741 ATTTCACACTACACACTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1800
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGAGTACCTGCGCGCGC 1860
Db 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGAGTACCTGCGCGCGC 1860
QY 1861 CATACTTTTGTAGTATTTCTGCGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1861 CATACTTTTGTAGTATTTCTGCGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
QY 1921 CAGTCTACTGCTGAGCTTTCAGCGCTTACGCGCTTACGAGTGAAGTGGTAAACTCGGAGTTG 1980
Db 1921 CAGTCTACTGCTGAGCTTTCAGCGCTTACGCGCTTACGAGTGAAGTGGTAAACTCGGAGTTG 1980
QY 1981 TAG 1983
Db 1981 TAG 1983

RESULT 5

US-08-542-634-1
; Sequence 1, Application US/08542634
; Patent No. 6214970
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; APPLICANT: McAtee, C. Patrick
; APPLICANT: Yarbough, Patrick O.
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/542.634
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO

; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
; INDIVIDUAL ISOLATE: ORF-2
; US-08-542-634-1

Query Match 100.0%; Score 1983; DB 3; Length 2049;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCGCCCTATTTTGTGTGCTCTCTCATGTTTTTGCCTATGCTGCCCGGCCA 60
Db 1 ATGCGCCCTCGCCCTATTTTGTGTGCTCTCTCATGTTTTTGCCTATGCTGCCCGGCCA 60
QY 61 CGCGCCGGTCAAGCGCTTGGCGCGCTGCTGGCGCGGAGCGGGTTCGGCGGTGT 120
Db 61 CGCGCCGGTCAAGCGCTTGGCGCGCTGCTGGCGCGGAGCGGGTTCGGCGGTGT 120
QY 121 TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 121 TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
QY 181 CCCTTGGCGCCCGATGTCAACGCTCGCGCGGGGCTGGACCTCTGTTTCGCCAACCCGCC 240
Db 181 CCCTTGGCGCCCGATGTCAACGCTCGCGCGGGGCTGGACCTCTGTTTCGCCAACCCGCC 240
QY 241 CGACCACTCGGCTCGGCTTGGCGTACAGGCGCCAGCGCCCGCGCTTCACGTGT 300
Db 241 CGACCACTCGGCTCGGCTTGGCGTACAGGCGCCAGCGCCCGCGCTTCACGTGT 300
QY 301 AGACCTACCAAGCTGGGGCGCGCGCTAAACCGGGTCTCCGCGCCCATGACACCCCG 360
Db 301 AGACCTACCAAGCTGGGGCGCGCGCTAAACCGGGTCTCCGCGCCCATGACACCCCG 360
QY 361 CCAGTGCCTGATGTGCACTCCCGCGCGCCCATCTTTGGCGCGCAGTATAACCTATCAACA 420
Db 361 CCAGTGCCTGATGTGCACTCCCGCGCGCCCATCTTTGGCGCGCAGTATAACCTATCAACA 420
QY 421 TCTCCCTTACCTCTTCGCTGGCGACCGGCACTAACCTGGTGTCTTTATGCGCGCCCTCT 480
Db 421 TCTCCCTTACCTCTTCGCTGGCGACCGGCACTAACCTGGTGTCTTTATGCGCGCCCTCT 480
QY 481 AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCCATATAATGGCCACGGAAGCTCT 540
Db 481 AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCCATATAATGGCCACGGAAGCTCT 540
QY 541 AATTATGCCAGTACCGGGTTCGCGTGCACCAATCCGTTACCGCGCGCTGTCCCAAT 600
Db 541 AATTATGCCAGTACCGGGTTCGCGTGCACCAATCCGTTACCGCGCGCTGTCCCAAT 600
QY 601 GCTGTGGCGGTAGCGCATCTCCATCTCTGCGCACAGACCAACACCCCGAGC 660
Db 601 GCTGTGGCGGTAGCGCATCTCCATCTCTGCGCACAGACCAACACCCCGAGC 660
QY 661 TCCGTTGATATGAATTCAAATAACCTCGACGGATGTTGTAATTTAGTCCAGCCGGCATA 720
Db 661 TCCGTTGATATGAATTCAAATAACCTCGACGGATGTTGTAATTTAGTCCAGCCGGCATA 720
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAAACCAAGGCTGGCGTCC 780
Db 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAAACCAAGGCTGGCGTCC 780
QY 781 GTGAGACCTCTGGGGTGGCTGAGGAGAGGCTACCTCTGCTCTTGTATGCTTTTGATA 840
Db 781 GTGAGACCTCTGGGGTGGCTGAGGAGAGGCTACCTCTGCTCTTGTATGCTTTTGATA 840
QY 841 CATGGCTCACTCGTAAATTCCTACTACTAATACACCTTACCGGTCCCTCGGGCTGTG 900
Db 841 CATGGCTCACTCGTAAATTCCTACTACTAATACACCTTACCGGTCCCTCGGGCTGTG 900
QY 901 GACCTTGCCTTGGAGCTTGAAGTTTCGCAACCTTACCCCCCGGTAAACCAATACCGGGTC 960
Db 901 GACCTTGCCTTGGAGCTTGAAGTTTCGCAACCTTACCCCCCGGTAAACCAATACCGGGTC 960

QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCCCTTCGTCGGGTGCGAGCGGACTCGCAG 1020
Db 961 TCCCGTTATTCAGCACTGCTCGCCACCGCCCTTCGTCGGGTGCGAGCGGACTCGCAG 1020
QY 1021 CTCACCAACAGGCTGTACCGCTTTATGAAGGACCTCTATTTTACTAGTAAATGTT 1080
Db 1021 CTCACCAACAGGCTGTACCGCTTTATGAAGGACCTCTATTTTACTAGTAAATGTT 1080
QY 1081 GTCGGTAGATCGCGCGCGGATAGCCCTCAACCTGTTCAACCTGCTGACACTCTGCTT 1140
Db 1081 GTCGGTAGATCGCGCGCGGATAGCCCTCAACCTGTTCAACCTGCTGACACTCTGCTT 1140
QY 1141 GCGGCTCGCAGACAGAAATGATTTCTCGCTGCTGGCGAGCTGTTCTACTCCGTCCT 1200
Db 1141 GCGGCTCGCAGACAGAAATGATTTCTCGCTGCTGGCGAGCTGTTCTACTCCGTCCT 1200
QY 1201 GTTGTCTCAGCAATGGCGAGCGCACTGTTAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db 1201 GTTGTCTCAGCAATGGCGAGCGCACTGTTAGTTGTATACATCTGTAGAGAATGCTCAG 1260
QY 1261 CAGGATAAGGTTATGCAATCCGCGATGACATTTGACTCGGAGAACTCGTGTGGTTAT 1320
Db 1261 CAGGATAAGGTTATGCAATCCGCGATGACATTTGACTCGGAGAACTCGTGTGGTTAT 1320
QY 1321 CAGGATTATCAACCAACATGACAGATCGCGCGAGCTTCTCAGCGCCCATCGCGC 1380
Db 1321 CAGGATTATGATAACCAACATGACAGATCGCGCGAGCTTCTCAGCGCCCATCGCGC 1380
QY 1381 CCTTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT 1440
Db 1381 CCTTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT 1440
QY 1441 GACAGTCACTATGCTCTTGAATGCGGCGAGTTATGTTTCTGACTCTGTGACTCTG 1500
Db 1441 GACAGTCACTATGCTCTTGAATGCGGCGAGTTATGTTTCTGACTCTGTGACTCTG 1500
QY 1501 GTTAATGTTGCGACCGCGCGAGCGCTTTCGCGGTGCTGATGACCAAGGTCA 1560
Db 1501 GTTAATGTTGCGACCGCGCGAGCGCTTTCGCGGTGCTGATGACCAAGGTCA 1560
QY 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCGCG 1620
Db 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCGCG 1620
QY 1621 CTCGCGGTAACTCTCTTTCTGGAGGAGGACCAACTAAAGCGGTTACCTTATAAT 1680
Db 1621 CTCGCGGTAACTCTCTTTCTGGAGGAGGACCAACTAAAGCGGTTACCTTATAAT 1680
QY 1681 TATAACCACTGCTAGCACTGCTGTTGAGAAATGCGCGCGGACCGGCTCGCT 1740
Db 1681 TATAACCACTGCTAGCACTGCTGTTGAGAAATGCGCGCGGACCGGCTCGCT 1740
QY 1741 ATTTCCACTTACCACTAGCTGCTGGTCTGGTCCGCTCTCAATTTCTGGGTTGCGGTT 1800
Db 1741 ATTTCCACTTACCACTAGCTGCTGGTCTGGTCCGCTCTCAATTTCTGGGTTGCGGTT 1800
QY 1801 TTAGCCCCCACTCTGCGTAGCAATGCTTGGAGTACCTTGGACTACCTGCGGCTTTC 1860
Db 1801 TTAGCCCCCACTCTGCGTAGCAATGCTTGGAGTACCTTGGACTACCTGCGGCTTTC 1860
QY 1861 CATACTTTTATGATTTCTGCGCAGAGTGGCGCCCTTGGCTTTCAGGGCTGCGGTTTC 1920
Db 1861 CATACTTTTATGATTTCTGCGCAGAGTGGCGCCCTTGGCTTTCAGGGCTGCGGTTTC 1920
QY 1921 CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTAAACTCTGGGAGTTG 1980
Db 1921 CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTAAACTCTGGGAGTTG 1980
QY 1981 TAG 1983
Db 1981 TAG 1983

RESULT 6
US-08-477-292-1
; Sequence 1, Application US/08477292
; Patent No. 6231641
; GENERAL INFORMATION:
; APPLICANT: Puerst, Thomas
; APPLICANT: McAttee, Patrick
; APPLICANT: Yarborough, Patrice
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cathleen M. Desjardins, M.D.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: US
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,292
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/327,952
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, Allan A
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: G32P5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)368-9500
; TELEFAX: (415)368-0709
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
; INDIVIDUAL ISOLATE: ORF-2
; US-08-477-292-1

Query Match 100.0%; Score 1983; DB 3; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTTTGCCTATGCTGCTGCGCGCCA 60
Db 1 ATGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTTTGCCTATGCTGCTGCGCGCCA 60
QY 61 CCGCGCGCTCAGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 120
Db 61 CCGCGCGCTCAGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 120
QY 121 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 121 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
QY 181 CCCTTGGCGCCCGATGTCACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 240
Db 181 CCCTTGGCGCCCGATGTCACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 240
QY 241 CGACCACTCGGCTCGGCTTGGCGTGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGT 300
Db 241 CGACCACTCGGCTCGGCTTGGCGTGACCGAGCGCGCGCGCGCGCGCGCGCGCGT 300

301 AGACCTACACAGCTGGGGCGCGCTAACCGGGTCCGCTCCGGCCCATGACACCCCG 360
Db |||||
301 AGACCTACACAGCTGGGGCGCGCTAACCGGGTCCGCTCCGGCCCATGACACCCCG 360
Qy |||||
361 CAGTGCCTGATGTCGACTCCCGCGGCGCATCTTTGGCGGCGAGTATAAAGCTATCAACA 420
Db |||||
361 CAGTGCCTGATGTCGACTCCCGCGGCGCATCTTTGGCGGCGAGTATAAAGCTATCAACA 420
Qy |||||
421 TCTCCCTTACCTCTTCGGTGGCGACCGGCACTAACCTGGTCTTTATGCGCGCTCTT 480
Db |||||
421 TCTCCCTTACCTCTTCGGTGGCGACCGGCACTAACCTGGTCTTTATGCGCGCTCTT 480
Qy |||||
481 AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATAATAGCCACGGAAGCTTCT 540
Db |||||
481 AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATAATAGCCACGGAAGCTTCT 540
Qy |||||
541 AATTATGCCAGTACCGGTTGCCGTCGACCAATCCGTTACCGCGCTGGTCCCAAT 600
Db |||||
541 AATTATGCCAGTACCGGTTGCCGTCGACCAATCCGTTACCGCGCTGGTCCCAAT 600
Qy |||||
601 GCTGTCGGGGTTAGCCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCGCGAG 660
Db |||||
601 GCTGTCGGGGTTAGCCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCGCGAG 660
Qy |||||
661 TCCGTTGATATGAATTCATTAACCTCGACGGATGTTCTGATTTTATGTCAGCGCCGCATA 720
Db |||||
661 TCCGTTGATATGAATTCATTAACCTCGACGGATGTTCTGATTTTATGTCAGCGCCGCATA 720
Qy |||||
721 GCCTCTGAGCTGTGATCCCAAGTGGGCTTACACTATCGTAACCAAGGCTGGGCTCC 780
Db |||||
721 GCCTCTGAGCTGTGATCCCAAGTGGGCTTACACTATCGTAACCAAGGCTGGGCTCC 780
Qy |||||
781 GTCGAGACTCTGGGGTGGCTGAGAGAGGCTACTCTGGTCTTGTATGCTTTGCGATA 840
Db |||||
781 GTCGAGACTCTGGGGTGGCTGAGAGAGGCTACTCTGGTCTTGTATGCTTTGCGATA 840
Qy |||||
841 CATGCTCACTCGTAAATTCCTATATACTAATACACCTATACCGGTGCCCTCGGGCTGTTG 900
Db |||||
841 CATGCTCACTCGTAAATTCCTATATACTAATACACCTATACCGGTGCCCTCGGGCTGTTG 900
Qy |||||
901 GACTTGGCCCTTGAGCTTTGAGTTTCGCAACCTTACCCCGGTAACCAATAACGCGGTC 960
Db |||||
901 GACTTGGCCCTTGAGCTTTGAGTTTCGCAACCTTACCCCGGTAACCAATAACGCGGTC 960
Qy |||||
961 TCCGTTATTCAGACTGCTCGCCACCGCTTCGTCGGGTGGCGGAGGACTGCCAG 1020
Db |||||
961 TCCGTTATTCAGACTGCTCGCCACCGCTTCGTCGGGTGGCGGAGGACTGCCAG 1020
Qy |||||
1021 CTCACCCAGCGCTGTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGGT 1080
Db |||||
1021 CTCACCCAGCGCTGTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGGT 1080
Qy |||||
1081 GTCGCTGAGTACGGCGCGGAGTACCTTCAACCTGTTTCAACCTTGTCTGACACTCTGCT 1140
Db |||||
1081 GTCGCTGAGTACGGCGCGGAGTACCTTCAACCTGTTTCAACCTTGTCTGACACTCTGCT 1140
Qy |||||
1141 GCGGCTGCGGACAGAAATGATTTCTGCTGGTGGTGGCCAGCTTCTACTCCCGTCC 1200
Db |||||
1141 GCGGCTGCGGACAGAAATGATTTCTGCTGGTGGTGGCCAGCTTCTACTCCCGTCC 1200
Qy |||||
1201 GTTGTCTCAGCAATGGCGAGCGGACTGTTAGTTGTATACATCTGTAGAGAACTCTCAG 1260
Db |||||
1201 GTTGTCTCAGCAATGGCGAGCGGACTGTTAGTTGTATACATCTGTAGAGAACTCTCAG 1260
Qy |||||
1261 CAGGATAAGGTTATGCAATCCCGATGACATTTGACCTCGGAGATCTCGTGTGTTATT 1320
Db |||||
1261 CAGGATAAGGTTATGCAATCCCGATGACATTTGACCTCGGAGATCTCGTGTGTTATT 1320
Qy |||||
1321 CAGGATTATGATAACCAATGAAAGATCGGCGGCGCTTCTCCAGCCCATCGCGC 1380
Db |||||
1321 CAGGATTATGATAACCAATGAAAGATCGGCGGCGCTTCTCCAGCCCATCGCGC 1380

1381 CCTTTCTCTGTCTTCAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCCCGAGTAT 1440
Db |||||
1381 CCTTTCTCTGTCTTCAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCCCGAGTAT 1440
Qy |||||
1441 GACCAGTCCACTTATGGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTGACTTG 1500
Db |||||
1441 GACCAGTCCACTTATGGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTGACTTG 1500
Qy |||||
1501 GTTATGTTGGACCGGCGCGAGCGGTCGCCGCTCGATTTGACAGGTCACA 1560
Db |||||
1501 GTTATGTTGGACCGGCGCGAGCGGTCGCCGCTCGATTTGACAGGTCACA 1560
Qy |||||
1561 CTGACGCTGGCCCTCTCCACCTCCAGCAGTACTCGAAGACCTTTCTTCTGCTGCGC 1620
Db |||||
1561 CTGACGCTGGCCCTCTCCACCTCCAGCAGTACTCGAAGACCTTTCTTCTGCTGCGC 1620
Qy |||||
1621 CTCGCGGTAAGCTCTCTTTCTGGAGGACGACCAACTAAAGCCGGGTACCTTATAAT 1680
Db |||||
1621 CTCGCGGTAAGCTCTCTTTCTGGAGGACGACCAACTAAAGCCGGGTACCTTATAAT 1680
Qy |||||
1681 TATACACACTGCTAGCGACCACTGCTTGTGAGAAATGCCGCGGACCGGTCGCT 1740
Db |||||
1681 TATACACACTGCTAGCGACCACTGCTTGTGAGAAATGCCGCGGACCGGTCGCT 1740
Qy |||||
1741 ATTTCCACTTACACCACTAGCTGGGTGCTGCTCCGCTCTCCATTTCTGCGGTTCCCGTT 1800
Db |||||
1741 ATTTCCACTTACACCACTAGCTGGGTGCTGCTCCGCTCTCCATTTCTGCGGTTCCCGTT 1800
Qy |||||
1801 TTAGCCCCCACTGCTGCTAGCAATGCTTGGAGTACTTTGGACTACCTTCCGCGCGC 1860
Db |||||
1801 TTAGCCCCCACTGCTGCTAGCAATGCTTGGAGTACTTTGGACTACCTTCCGCGCGC 1860
Qy |||||
1861 CATACTTTGATGATTTCTGCCAGAGTCCGCCCTTTGGCTTTCAGGCTGCGGTTTC 1920
Db |||||
1861 CATACTTTGATGATTTCTGCCAGAGTCCGCCCTTTGGCTTTCAGGCTGCGGTTTC 1920
Qy |||||
1921 CAGTCTACTGCTGCTGAGCTTCCAGCGCTTAAGATGAAGTGGTAAACTCGGAGTTG 1980
Db |||||
1921 CAGTCTACTGCTGCTGAGCTTCCAGCGCTTAAGATGAAGTGGTAAACTCGGAGTTG 1980
Qy |||||
1981 TAG 1983
Db |||||
1981 TAG 1983

RESULT 7
US-07-870-985A-3
; Sequence 3, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 7
; US-07-870-985A-3

Query Match      100.0%; Score 1983; DB 4; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGCGCCTCGGCTATTTGTTGCTGCTCCTCATGTTTGTGCTATGCTGCTGCGGCGCA 60
DB 1 ATCGCGCCTCGGCTATTTGTTGCTGCTCCTCATGTTTGTGCTATGCTGCTGCGGCGCA 60

QY 61 CCGCCCGGTCAGCGCTGCGCGCGGTCGTTGCGCGCGCGAGCGCGGTTCCGCGGTTGT 120
DB 61 CCGCCCGGTCAGCGCTGCGCGCGGTCGTTGCGCGCGCGAGCGCGGTTCCGCGGTTGT 120

QY 121 TTCTGGGGTGACCGGGTTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
DB 121 TTCTGGGGTGACCGGGTTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180

QY 181 CCCTTGCGCCCGCATGTACCGGTGCGCGCGGCGCTGGAATCGTTTCGCCAACCGGCC 240
DB 181 CCCTTGCGCCCGCATGTACCGGTGCGCGCGGCGCTGGAATCGTTTCGCCAACCGGCC 240

QY 241 CGACCACTCGGCTCGGCTTGGCGTGAACAGCGCGCGCGCGCGGTTGCTCAGCTCGT 300
DB 241 CGACCACTCGGCTCGGCTTGGCGTGAACAGCGCGCGCGCGCGGTTGCTCAGCTCGT 300

QY 301 AGACCTACCACTGCGGCGCGCGCGCTAACCGCGTCCGCGCCATGACACCCCG 360
DB 301 AGACCTACCACTGCGGCGCGCGCGCTAACCGCGTCCGCGCCATGACACCCCG 360

QY 361 CCAGTGCTGATGTGCACTCCGCGGCGCATCTTGGCGCGGAGTAAACCTATCAACA 420
DB 361 CCAGTGCTGATGTGCACTCCGCGGCGCATCTTGGCGCGGAGTAAACCTATCAACA 420

QY 421 TCTCCCTTACTCTTCCGTTGGCGACCGGCACTAACCTGGTTCTTTATGCGGCGCTCTT 480
DB 421 TCTCCCTTACTCTTCCGTTGGCGACCGGCACTAACCTGGTTCTTTATGCGGCGCTCTT 480

```

```

QY 481 AGTCCGCTTTTACCCCTTCAGGACGSCACCAATACCCATATATATGCGCACGGAAGCTTCT 540
DB 481 AGTCCGCTTTTACCCCTTCAGGACGSCACCAATACCCATATATATGCGCACGGAAGCTTCT 540

QY 541 AATTATGCCCAAGTACCGGTTGCGCGTGCCCAATCCGTTACCGCCGCTGGTCCCAAT 600
DB 541 AATTATGCCCAAGTACCGGTTGCGCGTGCCCAATCCGTTACCGCCGCTGGTCCCAAT 600

QY 601 GCTGTCGGGGTTAGCCATCTCCATCTCAITTTCTGSCACAGACACCAACCCCGACG 660
DB 601 GCTGTCGGGGTTAGCCATCTCCATCTCAITTTCTGSCACAGACACCAACCCCGACG 660

QY 661 TCGGTTGATATGAATTTCAATAAACCTCGACGGATGTTGCTATTTTAGTCCAGCCGCA 720
DB 661 TCGGTTGATATGAATTTCAATAAACCTCGACGGATGTTGCTATTTTAGTCCAGCCGCA 720

QY 721 GCCTCTGAGCTTTGTGATCCAAAGTGAGCGCCTACACTATCGTAACCAAGGCTGGCTCC 780
DB 721 GCCTCTGAGCTTTGTGATCCAAAGTGAGCGCCTACACTATCGTAACCAAGGCTGGCTCC 780

QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
DB 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840

QY 841 CATGGCTCACTCGTAAATTCCTATATAACACCTTACACCTATACCGTCCCTCGGGCTGTTG 900
DB 841 CATGGCTCACTCGTAAATTCCTATATAACACCTTACACCTATACCGTCCCTCGGGCTGTTG 900

QY 901 GACTTTGCCCTTGAGCTTCAGTTTGCACCTTACCCCGGTAAACCAATACCGCGGTC 960
DB 901 GACTTTGCCCTTGAGCTTCAGTTTGCACCTTACCCCGGTAAACCAATACCGCGGTC 960

QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTCGGGTGCGGAGCGGAGTCCGCG 1020
DB 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTCGGGTGCGGAGCGGAGTCCGCG 1020

QY 1021 CTCACCAACGCGCTGCTACCCGCTTTATGAAGCACTCTATTTTACTAGTAAATGGT 1080
DB 1021 CTCACCAACGCGCTGCTACCCGCTTTATGAAGCACTCTATTTTACTAGTAAATGGT 1080

QY 1081 GTCGGTGAATCGGCGCGGATAGCCCTCACCTGTTCAACCTGTTGACATCTGCTGCTT 1140
DB 1081 GTCGGTGAATCGGCGCGGATAGCCCTCACCTGTTCAACCTGTTGACATCTGCTGCTT 1140

QY 1141 GCGCGCTCGCGACAAATTTGATTTGCTGCGGTGGGCGAGCTGTTTACTCCGTC 1200
DB 1141 GCGCGCTCGCGACAAATTTGATTTGCTGCGGTGGGCGAGCTGTTTACTCCGTC 1200

QY 1201 GTTGTCTCAGCCAAATGGCGAGCGCACTGTTAAAGTTGTATATCATCTGTAGAGAATGCTCAG 1260
DB 1201 GTTGTCTCAGCCAAATGGCGAGCGCACTGTTAAAGTTGTATATCATCTGTAGAGAATGCTCAG 1260

QY 1261 CAGGATAAGGGTATTGCAATTCGCGATGACATTTGAGTCTGGAGAAATCTCGTGGTTATT 1320
DB 1261 CAGGATAAGGGTATTGCAATTCGCGATGACATTTGAGTCTGGAGAAATCTCGTGGTTATT 1320

QY 1321 CAGGATTATGATAACCAACATGAACAGATCGCGCGAGCTTCTCAGCCCATCGCG 1380
DB 1321 CAGGATTATGATAACCAACATGAACAGATCGCGCGAGCTTCTCAGCCCATCGCG 1380

QY 1381 CCTTTCTGCTCTCGAGCTAATGATGTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
DB 1381 CCTTTCTGCTCTCGAGCTAATGATGTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440

QY 1441 GACAGTCACTATGCTCTTACCTGCGGCGAGTTATGTTTCTGACTGTTGAGCTTGG 1500
DB 1441 GACAGTCACTATGCTCTTACCTGCGGCGAGTTATGTTTCTGACTGTTGAGCTTGG 1500

QY 1501 GTTAAATGTCGACCGCGCGAGCGGCTTCCCGGTGCGTTCGATTGGAACCAAGGTCA 1560
DB 1501 GTTAAATGTCGACCGCGCGAGCGGCTTCCCGGTGCGTTCGATTGGAACCAAGGTCA 1560

```

1561 CTTGACGGTGGCCCTCTCCACATCCAGAGTACTCGAAGACCTCTTTGTCTGCGG 1620
Db CTTGACGGTGGCCCTCTCCACATCCAGAGTACTCGAAGACCTCTTTGTCTGCGG 1620
1621 CTCGCGGTAAGCTCTCTTTCTGGGAGGAGGACACAACTAAAGCGGGTACCTTATAAT 1680
Db CTCGCGGTAAGCTCTCTTTCTGGGAGGAGGACACAACTAAAGCGGGTACCTTATAAT 1680
1681 TATAACACCACTCTAGGACCACTGCTTGTGAGAAATCCCGCGGACCGGGTGGCT 1740
Db TATAACACCACTCTAGGACCACTGCTTGTGAGAAATCCCGCGGACCGGGTGGCT 1740
1741 ATTTCACCTTACACCACTAGCTGGGTGGTGGTCCGCTCTCCATTTCTGCGGTTGCCGTT 1800
Db ATTTCACCTTACACCACTAGCTGGGTGGTGGTCCGCTCTCCATTTCTGCGGTTGCCGTT 1800
1801 TTAGCCCCCACTCTGGCTAGCAATGCTTGGAGATACCTTGAGCTACCTTGCGCGGC 1860
Db TTAGCCCCCACTCTGGCTAGCAATGCTTGGAGATACCTTGAGCTACCTTGCGCGGC 1860
1861 CATACCTTTTGTGATGATTTCTGCCCCAGAGTGGCGGCCCTTGGCCTTCAGGCTGCGCTTTC 1920
Db CATACCTTTTGTGATGATTTCTGCCCCAGAGTGGCGGCCCTTGGCCTTCAGGCTGCGCTTTC 1920
1921 CAGTCTACTGCTGCTGAGCTTCAGCCCTTAAGATGAAGTGGTAAACTCGGGAGTTG 1980
Db CAGTCTACTGCTGCTGAGCTTCAGCCCTTAAGATGAAGTGGTAAACTCGGGAGTTG 1980
1981 TAG 1983
Db TAG 1983

RESULT 8

PCT-US95-13703-1
; Sequence 1, Application PC/TUS9513703
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
; TITLE OF INVENTION: US95 THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Deklinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13703
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
; INDIVIDUAL ISOLATE: ORF-2
; PCT-US95-13703-1
Query Match 100.0%; Score 1983; DB 5; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGTCCTGCGGCCA 60
Db 1 ATGCGGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGTCCTGCGGCCA 60
QY 61 CGCGCCGGTCTAGCGCTCTGCGCGCTGCTGGGCGGCGAGCGCGGTTCGCGCGGTGT 120
Db 61 CGCGCCGGTCTAGCGCTCTGCGCGCTGCTGGGCGGCGAGCGCGGTTCGCGCGGTGT 120
QY 121 TTCTGCGGCTGACCGGGTTCGATTTCTCAGCCCTTCGCAATCCCTTATATTCATCCAAAC 180
Db 121 TTCTGCGGCTGACCGGGTTCGATTTCTCAGCCCTTCGCAATCCCTTATATTCATCCAAAC 180
QY 181 CCGTTGCGCCCGATGTCACCGCTGCGCGCGGGCTGGACCTGCTGTCGCAACCCGCC 240
Db 181 CCGTTGCGCCCGATGTCACCGCTGCGCGCGGGCTGGACCTGCTGTCGCAACCCGCC 240
QY 241 CGACCACTCGGCTCGGCTTGGCTGACAGGCGCGCGCGCGCGCTGCTGCTGCTGCT 300
Db 241 CGACCACTCGGCTCGGCTTGGCTGACAGGCGCGCGCGCGCGCTGCTGCTGCTGCT 300
QY 301 AGACCTTACCACTGCGGCGCGCGCTTAAACCGCGCTGCTCGGCGCGCATGACACCCG 360
Db 301 AGACCTTACCACTGCGGCGCGCGCTTAAACCGCGCTGCTCGGCGCGCATGACACCCG 360
QY 361 CCAGTCTGCTGATGTCGACTCCCGCGCGCGCATCTTTGCGCGCGAGTATAAAGCTATCA 420
Db 361 CCAGTCTGCTGATGTCGACTCCCGCGCGCGCATCTTTGCGCGCGAGTATAAAGCTATCA 420
QY 421 TCTCCCTTACCTCTTCGCTGCGCGCGCGCATCTTTGCGCGCGAGTATAAAGCTATCA 480
Db 421 TCTCCCTTACCTCTTCGCTGCGCGCGCGCATCTTTGCGCGCGAGTATAAAGCTATCA 480
QY 481 AGTCCGCTTTTACCCCTTTCAGGACCGGCAATACCCATATATTAATGCGCGAGGCTCT 540
Db 481 AGTCCGCTTTTACCCCTTTCAGGACCGGCAATACCCATATATTAATGCGCGAGGCTCT 540
QY 541 AATTATGCCAGTACCGGCTTGGCGTGCACCAATCCGTTACCGCGCGGTGCTGCTGCT 600
Db 541 AATTATGCCAGTACCGGCTTGGCGTGCACCAATCCGTTACCGCGCGGTGCTGCTGCT 600
QY 601 GCTGCGCGGTTAGCCATCTCCATCTTCATCTGCGCGCGAGACACCAACCGCGAG 660
Db 601 GCTGCGCGGTTAGCCATCTTCATCTTCATCTGCGCGCGAGACACCAACCGCGAG 660
QY 661 TCCGTTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
Db 661 TCCGTTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db 721 GCCTCTGAGCTTGTGATCCCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 781 GTCGAGACCTCTGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db 781 GTCGAGACCTCTGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
QY 841 CATGGCTCACTCGTAAATTCCTTACTAATACCCCTTACCTATCGTAAACCAAGGCTGG 900
Db 841 CATGGCTCACTCGTAAATTCCTTACTAATACCCCTTACCTATCGTAAACCAAGGCTGG 900
QY 901 GACTTTGCGCTTGTGAGCTTGTGAGTTCGCAACCTTACCCCGGTAACCAATACCGGGTC 960
Db 901 GACTTTGCGCTTGTGAGCTTGTGAGTTCGCAACCTTACCCCGGTAACCAATACCGGGTC 960
QY 961 TCCCGTTATTCAGCACTGCTGCGCGCATCGGCTTCTGCTGCGGTCGGAGGAGTCTGCG 1020

QY 61 CGCCCGGTCAGCGCTCTGGCGCGCGTCTGGTGGCGCGCGAGCGCGGTTCGCGCGTGGT 120
DB |||||
QY 106 CGCCCGGTCAGCGCTCTGGCGCGCGTCTGGTGGCGCGCGAGCGCGGTTCGCGCGTGGT 165
DB |||||
QY 121 TTCGGGTGACCGGGTGGTGAATCTCAGCGCTTCGCAATCCCGCTATATTCATCAACCAAC 180
DB |||||
QY 166 TTCGGGTGACCGGGTGGTGAATCTCAGCGCTTCGCAATCCCGCTATATTCATCAACCAAC 225
DB |||||
QY 181 CCCTTGGCCCCGATGTACCGCTGGGCGGGGCTGGACCTCTGTGTTCGCAACCCCGC 240
DB |||||
QY 241 CGACCACTCGGCTCGGCTTGGGTGACGAGCGCCAGCGCCCGCGCTTGGCTCAAGTGGT 300
DB |||||
QY 286 CGACCACTCGGCTCGGCTTGGGTGACGAGCGCCAGCGCCCGCGCTTGGCTCAAGTGGT 345
DB |||||
QY 301 AGACCTACACAGCTGGGGCGCGCGCTTAACCGCGGTGCTTCGCGGCCCATGACACCCCG 360
DB |||||
QY 346 AGACCTACACAGCTGGGGCGCGCGCTTAACCGCGGTGCTTCGCGGCCCATGACACCCCG 405
DB |||||
QY 361 CGAGTGCCTGATGTCGACTCCGCGGCGGCATCTTTGGCGCGGAGTAACTATCAACA 420
DB |||||
QY 406 CCAGTGCCTGATGTCGACTCCGCGGCGGCATCTTTGGCGCGGAGTAACTATCAACA 465
DB |||||
QY 421 TCTCCCTTACCTCTTCCGTGGCCACCGGCACTAACCTGGTCTTTATGCGCCCTCTT 480
DB |||||
QY 466 TCTCCCTTACCTCTTCCGTGGCCACCGGCACTAACCTGGTCTTTATGCGCCCTCTT 525
DB |||||
QY 481 AGTCGGCTTTACCCCTTCAGGACGGCACCAATACCCATATAATGGCCACGGAAGCTTCT 540
DB |||||
QY 526 AGTCGGCTTTACCCCTTCAGGACGGCACCAATACCCATATAATGGCCACGGAAGCTTCT 585
DB |||||
QY 541 AATATGCCAGTACCGGGTGGCGGTCGACCAATTCGGTTACCGCGCGTGGTCCCAAT 600
DB |||||
QY 586 AATATGCCAGTACCGGGTGGCGGTCGACCAATTCGGTTACCGCGCGTGGTCCCAAT 645
DB |||||
QY 601 GCTGTCGGGGTACCGCATCTCCATCTCATCTGCGCCACAGACCAACCAACCCGAG 660
DB |||||
QY 646 GCTGTCGGGGTACCGCATCTCCATCTCATCTGCGCCACAGACCAACCAACCCGAG 705
DB |||||
QY 661 TCCGTTGATATGAATTCATTAACCTCGAGCGATGTTGGTATTTTATGTCAGCGCGCAT 720
DB |||||
QY 706 TCCGTTGATATGAATTCATTAACCTCGAGCGATGTTGGTATTTTATGTCAGCGCGCAT 765
DB |||||
QY 721 GCCTCTGAGCTTGATCCCAAGTGAAGCGCTACACTATCGTAAACAGGCTGGCGCTCC 780
DB |||||
QY 766 GCCTCTGAGCTTGATCCCAAGTGAAGCGCTACACTATCGTAAACAGGCTGGCGCTCC 825
DB |||||
QY 781 GTCGAGACTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTGTTTATGCTTTGCATA 840
DB |||||
QY 826 GTCGAGACTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTGTTTATGCTTTGCATA 885
DB |||||
QY 841 CATGCTCACTCGTAAATTCCTATTAATAACCTATACACCTATACCGGTGCGCTCGGGCTGTTG 900
DB |||||
QY 886 CATGCTCACTCGTAAATTCCTATTAATAACCTATACACCTATACCGGTGCGCTCGGGCTGTTG 945
DB |||||
QY 901 GACTTGGCTTGGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGTC 960
DB |||||
QY 946 GACTTGGCTTGGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGTC 1005
DB |||||
QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCTGTCGGGTGGGAGGAGTCCCGAG 1020
DB |||||
QY 1006 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCTGTCGGGTGGGAGGAGTCCCGAG 1065
DB |||||
QY 1021 CTCACCAACCGCTGCTACCGCTTTTATGAAGGACCTCTATTTTACTAGTAACTAATGTT 1080
DB |||||
QY 1066 CTCACCAACCGCTGCTACCGCTTTTATGAAGGACCTCTATTTTACTAGTAACTAATGTT 1125
DB |||||
QY 1081 GTCGTTGAGTACGCGCGGGAGTACCGCTTCAACCTGTTCAACCTGCTGACACTCTGTT 1140
DB |||||
QY 1126 GTCGTTGAGTACGCGCGGGAGTACCGCTTCAACCTGTTCAACCTGCTGACACTCTGTT 1185
DB |||||

QY 1141 GGCGCGCTGCGCAGAGAAATGATTTCTGTCGGCTGGTGGCGAGCTGTTCTACTCCCGTCC 1200
DB |||||
QY 1186 GGCGCGCTGCGCAGAGAAATGATTTCTGTCGGCTGGTGGCGAGCTGTTCTACTCCCGTCC 1245
DB |||||
QY 1201 GTTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATATACATCTGTAGAGAAATGCTCAG 1260
DB |||||
QY 1246 GTTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATATACATCTGTAGAGAAATGCTCAG 1305
DB |||||
QY 1261 CAGGTAAGGTATTTGCAATCCCGCATGACATGACCTCGGAGAAATCTCGTGTGGTATT 1320
DB |||||
QY 1306 CAGGTAAGGTATTTGCAATCCCGCATGACATGACCTCGGAGAAATCTCGTGTGGTATT 1365
DB |||||
QY 1321 CAGGTAATGATAACCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1380
DB |||||
QY 1366 CAGGTAATGATAACCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1425
DB |||||
QY 1381 CCTTTCTCTGCTCTTTCGAGCTAAATGATGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
DB |||||
QY 1426 CCTTTCTCTGCTCTTTCGAGCTAAATGATGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1485
DB |||||
QY 1441 GACAGTCCACTATATGCTCTTCGACTGCGCGGAGTATTTATGTTTCTGACTCTGTGACCTTG 1500
DB |||||
QY 1486 GACAGTCCACTATATGCTCTTCGACTGCGCGGAGTATTTATGTTTCTGACTCTGTGACCTTG 1545
DB |||||
QY 1501 GTTAATGTTGGACCGCGCGGAGCGGCTGCGCGGCTGCTGATTTGGACCAAGGTCA 1560
DB |||||
QY 1546 GTTAATGTTGGACCGCGCGGAGCGGCTGCGCGGCTGCTGATTTGGACCAAGGTCA 1605
DB |||||
QY 1561 CTGACGGTTCGCGCGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCTCGCG 1620
DB |||||
QY 1606 CTGACGGTTCGCGCGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCTCGCG 1665
DB |||||
QY 1621 CTCGCGGTAAAGTCTCTTTCTGGAGGAGCAGCAAACTAAAGCCGGGTACCTTATTAAT 1680
DB |||||
QY 1666 CTCGCGGTAAAGTCTCTTTCTGGAGGAGCAGCAAACTAAAGCCGGGTACCTTATTAAT 1725
DB |||||
QY 1681 TATAACACCTGCTAGCAGCAAACTGTCGAGAAATGCGCGCGGAGCAGCGGTCGCT 1740
DB |||||
QY 1726 TATAACACCTGCTAGCAGCAAACTGTCGAGAAATGCGCGCGGAGCAGCGGTCGCT 1785
DB |||||
QY 1741 ATTTCACTTACACCATAGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB |||||
QY 1786 ATTTCACTTACACCATAGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1845
DB |||||
QY 1801 TTAGCCCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB |||||
QY 1846 TTAGCCCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1905
DB |||||
QY 1861 CATACTTTTGTATGATTTCTGCGCGAGTGGCGCGCTTGGCTTTCAGGGCTGCGCTTTC 1920
DB |||||
QY 1906 CATACTTTTGTATGATTTCTGCGCGAGTGGCGCGCTTGGCTTTCAGGGCTGCGCTTTC 1965
DB |||||
QY 1921 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1980
DB |||||
QY 1966 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2025
DB |||||
QY 1981 TAG 1983
DB 2026 TAG 2028

RESULT 10

US-08-484-054-1

; Sequence 1, Application US/08484054

; Patent No. 5770689

; GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory R.

; APPLICANT: Bradley, Daniel W.

; APPLICANT: Twd, Jr-Shin

; APPLICANT: Purdy, Michael A.

; APPLICANT: Tam, Albert W.

; APPLICANT: Krawczynski, Krzysztof Z.

; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method

NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,054
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 870,985
FILING DATE: 20-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.38
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2094 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: BURMA SEQUENCE, FIGURE 7

US-08-484-054-1
Query Match 100.0%; Score 1983; DB 1; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCGCGCCCTGGGCTATTTGTTGCTGCTCATGTTTGTGCTATGCTGCTGCGCGCCA 60
Db 46 ATCGCGCCCTGGGCTATTTGTTGCTGCTCATGTTTGTGCTATGCTGCTGCGCGCCA 105
QY 61 CCGCGCGCTCAGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 106 CCGCGCGCTCAGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 165
QY 121 TTCTGCGGGTACCGGGTGTGTTCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC 180
Db 166 TTCTGCGGGTACCGGGTGTGTTCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC 225

181 CCCTTCGCCCCCGAATGTACCGCTGCGCGCGGGCTGGACCTCGTGTTCGCCAACCGGCC 240
226 CCCTTCGCCCCCGAATGTACCGCTGCGCGCGGGCTGGACCTCGTGTTCGCCAACCGGCC 285
241 CGACCACTCGGCTCCGCTTGGGGTGACAGGGCCAGCGCCCGCGCTGCGCTCACGTCGT 300
286 CGACCACTCGGCTCCGCTTGGGGTGACAGGGCCAGCGCCCGCGCTGCGCTCACGTCGT 345
301 AGACCTACCACTGCGGCGCGCGCTAACCGCGCTCGCTCCGCGCCCATGACACCCCG 360
346 AGACCTACCACTGCGGCGCGCGCGCTAACCGCGCTCGCTCCGCGCCCATGACACCCCG 405
361 CCAGTGCCTGATGTGACTCCGCGCGCGCATCTTCGCGCGGCGAGTAACTATCAACA 420
406 CCAGTGCCTGATGTGACTCCGCGCGCGCATCTTCGCGCGGCGAGTAACTATCAACA 465
421 TCTCCCTTACCTTTCCTGCGCGCGCATCTAACCTGGTTCCTTTATGCGCGCCCTCTT 480
466 TCTCCCTTACCTTTCCTGCGCGCATCTAACCTGGTTCCTTTATGCGCGCCCTCTT 525
481 AGTCCGCTTTTACCCCTTCAGGACGCGCAATACCATATATATGCGCGAGGCTTCT 540
526 AGTCCGCTTTTACCCCTTCAGGACGCGCAATACCATATATATGCGCGAGGCTTCT 585
541 AATTATGCCAGTACCGGGTTGCGCGTGCCACAATCCGTTACCGCGCGTGTGCCCAAT 600
586 AATTATGCCAGTACCGGGTTGCGCGTGCCACAATCCGTTACCGCGCGTGTGCCCAAT 645
601 GCTGCGCGGGTTACCGCATCTCCATCTCATCTGCGCGACAGACACACCGCGCGAG 660
646 GCTGCGCGGGTTACCGCATCTCCATCTCATCTGCGCGACAGACACACCGCGCGAG 705
661 TCCGTTGATGAAATTCATTAACCTCGACGGATGTTGCTATTTAGTCCAGCCCGCAT 720
706 TCCGTTGATGAAATTCATTAACCTCGACGGATGTTGCTATTTAGTCCAGCCCGCAT 765
721 GCCTCTGAGCTTTGATCCCAAGTGAGCGCCTACATATGTAACCAAGCTGGCGCTCC 780
766 GCCTCTGAGCTTTGATCCCAAGTGAGCGCCTACATATGTAACCAAGCTGGCGCTCC 825
781 GTCGAGACCTCTGGGGTGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
826 GTCGAGACCTCTGGGGTGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 885
841 CATGGCTCACTGTAATTTCTTACTATAFACACCTATACCGGTGCGCTCGGGCTGTTG 900
886 CATGGCTCACTGTAATTTCTTACTATAFACACCTATACCGGTGCGCTCGGGCTGTTG 945
901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATACGCGGTC 960
946 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATACGCGGTC 1005
961 TCCGTTATTCAGCACTGCTGCGCACCGCTTCGTGCGGGTGGAGCGGAGTCCGCG 1020
1006 TCCGTTATTCAGCACTGCTGCGCACCGCTTCGTGCGGGTGGAGCGGAGTCCGCG 1065
1021 CTCACCAACGAGCTGCTACCGCTTTATGAAGAGCTCTATTTTACTAGTACTAATGTT 1080
1066 CTCACCAACGAGCTGCTACCGCTTTATGAAGAGCTCTATTTTACTAGTACTAATGTT 1125
1081 GTCGTTGAGATCGCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTGCTT 1140
1126 GTCGTTGAGATCGCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTGCTT 1185
1141 GCGCGCTCGCGAGCAAGTATGTTTGTGCGGGTGGCGAGCTGTTCTACTCCGTC 1200
1186 GCGCGCTCGCGAGCAAGTATGTTTGTGCGGGTGGCGAGCTGTTCTACTCCGTC 1245
1201 GTTGTCTCAGCAATGCGGAGCGGACTGTTTAACTTGTATACATCTGTAGAGATGCTCAG 1260
1246 GTTGTCTCAGCAATGCGGAGCGGACTGTTTAACTTGTATACATCTGTAGAGATGCTCAG 1305
1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTCGAGAACTCTCGTGGGTATT 1320

```
Db 1306 CAGGATAAGGGTATTGGAATCCCGCATGACATTGACCTCGGAGATCTCGTGTGTTATT 1365
Qy 1321 CAGGATTATGATAACCAACATGAACAGATCGGCGAGCGCTTCTCCAGCCCGCATCGCGC 1380
Db 1366 CAGGATTATGATAACCAACATGAACAGATCGGCGAGCGCTTCTCCAGCCCGCATCGCGC 1425
Qy 1381 CCTTCTCTGCTGCTTTCGAGCTTAATGATGCTTTTGGCTCTCTCTCACCGCTCCGAGTAT 1440
Db 1426 CCTTCTCTGCTGCTTTCGAGCTTAATGATGCTTTTGGCTCTCTCTCACCGCTCCGAGTAT 1485
Qy 1441 GACCAGTCCACTTATGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 1486 GACCAGTCCACTTATGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1545
Qy 1501 GTTAATGTTGGACCGGCGGCGAGCGGCTTGGCCCGGTGCTCGATGACCAAGGTCA 1560
Db 1546 GTTAATGTTGGACCGGCGGCGAGCGGCTTGGCCCGGTGCTCGATGACCAAGGTCA 1605
Qy 1561 CTTGACGGTTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCTGCG 1620
Db 1606 CTTGACGGTTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCTGCG 1665
Qy 1621 CTCGCGGTAAAGTCTCTTTCTGGAGGAGGAGGACCAACTAAAGCGGGTACCTTTATAT 1680
Db 1666 CTCGCGGTAAAGTCTCTTTCTGGAGGAGGAGGACCAACTAAAGCGGGTACCTTTATAT 1725
Qy 1681 TATAACACACTGTAGGACCAACTGCTTGTGAGATGCGCGGGGACCGGTGCT 1740
Db 1726 TATAACACACTGTAGGACCAACTGCTTGTGAGATGCGCGGGGACCGGTGCT 1785
Qy 1741 ATTTCACCTTACACACTAGCTGGGTGCTGCTCCGCTCCATTTCTGCGGTGCGCTT 1800
Db 1786 ATTTCACCTTACACACTAGCTGGGTGCTGCTCCGCTCCATTTCTGCGGTGCGCTT 1845
Qy 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGGAGTACTTGGACTACCTGCCCGCGC 1860
Db 1846 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGGAGTACTTGGACTACCTGCCCGCGC 1905
Qy 1861 CATACCTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCTTTCAGGGTTCGCTTTC 1920
Db 1906 CATACCTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCTTTCAGGGTTCGCTTTC 1965
Qy 1921 CAGTCTACTGCTGCTGAGCTTACAGCGCTTAAGATGAAGTGGTAAACTCGGAGTTG 1980
Db 1966 CAGTCTACTGCTGCTGAGCTTACAGCGCTTAAGATGAAGTGGTAAACTCGGAGTTG 2025
Qy 1981 TAG 1983
Db 2026 TAG 2028
```

```
RESULT 11
US-07-876-941A-1
; Sequence 1, Application US/07876941A
; Patent No. 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; TITLE OF INVENTION: Antibodies
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/876.941A
; FILING DATE: 01-MAY-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.33
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2094 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: BURMA SEQUENCE, FIGURE 7
US-07-876-941A-1
```

```
Query Match 100.0%; Score 1983; DB 2; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGCCCTCGGCGCTATTGTTGCTGCTCTCATGTTTTCCTATGCTGCGCGCGCA 60
Db 46 ATGGCGCCCTCGGCGCTATTGTTGCTGCTCTCATGTTTTCCTATGCTGCGCGCGCA 105
Qy 61 CCGCCCGGTTCAGCGCTCTGGCGCGCGCTGCTGGCGCGCGAGCGGTTCCGCGCGGTG 120
Db 106 CCGCCCGGTTCAGCGCTCTGGCGCGCGCTGCTGGCGCGCGAGCGGTTCCGCGCGGTG 165
Qy 121 TTCTGGGTGACCGGTTGATTTCTCAGCCCTTCGCAATCCCTATATATTCACCAAC 180
Db 166 TTCTGGGTGACCGGTTGATTTCTCAGCCCTTCGCAATCCCTATATATTCACCAAC 225
Qy 181 CCCTTTCGCCCGCGGATGTCAACCGCTGCGCGCGCGCTGAGCTGTTGCGCAACCGCC 240
Db 226 CCCTTTCGCCCGCGGATGTCAACCGCTGCGCGCGCGCTGAGCTGTTGCGCAACCGCC 285
Qy 241 CGACACTTCGGCTCGGCTTGGCTGACCAAGCGCGCGCGCGCGCTTTCCTCAGCTG 300
Db 286 CGACACTTCGGCTCGGCTTGGCTGACCAAGCGCGCGCGCGCTTTCCTCAGCTG 345
Qy 301 AGACTTACCAAGCTGGGGCGCGCGCTAAACCGCGGTGCTTCGCGCGCGCATGACAC 360
Db 346 AGACTTACCAAGCTGGGGCGCGCGCTAAACCGCGGTGCTTCGCGCGCGCATGACAC 405
Qy 361 CCAGTGCCTGATGTCGACTCCCGCGCGCGCATCTTTCGCGCGCGAGTATAACCTATCA 420
```


Db 406 CAGTGGCTGATGTGACATCCCGGGGCACTTTCGGCGGAGTATACCTATCAACA 465
QY 421 TCTCCCTTAACTCTTTCCTGGGCAACCGGACATACTGGTCTCTTTATGCGGCCCTCTT 480
Db 466 TCTCCCTTAACTCTTTCCTGGGCAACCGGACATACTGGTCTCTTTATGCGGCCCTCTT 525
QY 481 AGTCGGCTTTTACCCCTTCAGAGCGGACCAATACCATATATATGAGCCAGAGCTTCT 540
Db 526 AGTCGGCTTTTACCCCTTCAGAGCGGACCAATACCATATATATGAGCCAGAGCTTCT 585
QY 541 AATTATGCCAGTACCGGGTTCGCCGTGCAAAATCCGTTTACCGCCCGCTGGTCCCAAT 600
Db 586 AATTATGCCAGTACCGGGTTCGCCGTGCAAAATCCGTTTACCGCCCGCTGGTCCCAAT 645
QY 601 GGTGTCGGGGTTACGCCATCTCCATCTCTATCTGCGGCAAGACCAACCGGCGAG 660
Db 646 GGTGTCGGGGTTACGCCATCTCCATCTCTATCTGCGGCAAGACCAACCGGCGAG 705
QY 661 TCCGTTGATATGAATTCATTAACCTCGAGGATGTTGCTATTTTATGTCAGCCCGGCATA 720
Db 706 TCCGTTGATATGAATTCATTAACCTCGAGGATGTTGCTATTTTATGTCAGCCCGGCATA 765
QY 721 GCCTCTGAGCTTGATGCCAAGTGAGCGGCTACACTATCGTAACCAAGGCTGGCGCTCC 780
Db 766 GCCTCTGAGCTTGATGCCAAGTGAGCGGCTACACTATCGTAACCAAGGCTGGCGCTCC 825
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGCTTGTATGCTTTGCATA 840
Db 826 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGCTTGTATGCTTTGCATA 885
QY 841 CATGGCTCACTGTAATTCCTATCTATATACACCTATACCGGTCGCCCTCGGGCTGTTG 900
Db 886 CATGGCTCACTGTAATTCCTATCTATATACACCTATACCGGTCGCCCTCGGGCTGTTG 945
QY 901 GACTTTGGCCCTTGAGCTTGAGTTTCGCAACTTACCCCGGTAACCAATACCGGGGTC 960
Db 946 GACTTTGGCCCTTGAGCTTGAGTTTCGCAACTTACCCCGGTAACCAATACCGGGGTC 1005
QY 961 TCCCGTTATTCAGACCTGTCGCGGACCGGCTTCGTCGGGTCGGAGCGGAGTCCGCGAG 1020
Db 1006 TCCCGTTATTCAGACCTGTCGCGGACCGGCTTCGTCGGGTCGGAGCGGAGTCCGCGAG 1065
QY 1021 CTCACACACAGCTGCTACCGCTTTATGAAGGACTCTATTTTACTAGTACTAATGTT 1080
Db 1066 CTCACACACAGCTGCTACCGCTTTATGAAGGACTCTATTTTACTAGTACTAATGTT 1125
QY 1081 GTCGGTGAGATCGGCGGGGATAGCCCTCACCTGTTCAACCTTGTGACACCTCTGCTT 1140
Db 1126 GTCGGTGAGATCGGCGGGGATAGCCCTCACCTGTTCAACCTTGTGACACCTCTGCTT 1185
QY 1141 GCGGGCTGCGGACAGATGATTTGTCGGCTGGTGGCCAGCTGTTCTACTCCGCTCC 1200
Db 1186 GCGGGCTGCGGACAGATGATTTGTCGGCTGGTGGCCAGCTGTTCTACTCCGCTCC 1245
QY 1201 GTTGTCTCAGCAATGCGGACGAGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db 1246 GTTGTCTCAGCAATGCGGACGAGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1305
QY 1261 CAGGATAAGGGTATGCAATCCCGCATGACATTAACCTCGGAGAACTCTCGTGGTATT 1320
Db 1306 CAGGATAAGGGTATGCAATCCCGCATGACATTAACCTCGGAGAACTCTCGTGGTATT 1365
QY 1321 CAGGATATGATAACCAAGATGACAGATGCGGCGAGCTTCTCAGCCCATCGGC 1380
Db 1366 CAGGATATGATAACCAAGATGACAGATGCGGCGAGCTTCTCAGCCCATCGGC 1425
QY 1381 CCTTCTCTCTCTCGAGCTTAATGATGCTTGGCTCTCTCTCAGCGCTGCGGAGTAT 1440
Db 1426 CCTTCTCTCTCTCGAGCTTAATGATGCTTGGCTCTCTCTCAGCGCTGCGGAGTAT 1485
QY 1441 GACCACTTCACTTATGGCTCTTGCAGCTGGGCCAGATTTATGTTCTGACTCTGAGCTTG 1500

Db 1486 GACCACTTCACTTATGGCTCTTTCGACTGGCCCAAGTTTATGTTTCTGACTCTGTGACCTTG 1545
QY 1501 GTTAAATGTTGCGACCGGCGGAGCGGCTTGGCCGCTCGATTTGGACCAAGGTCA 1560
Db 1546 GTTAAATGTTGCGACCGGCGGAGCGGCTTGGCCGCTCGATTTGGACCAAGGTCA 1605
QY 1561 CTTGACGGTGGCGGCCCTCTCTCCACCATCCAGAGTACTCGAAAGACCTTCTTTGTCCTGCGG 1620
Db 1606 CTTGACGGTGGCGGCCCTCTCTCCACCATCCAGAGTACTCGAAAGACCTTCTTTGTCCTGCGG 1665
QY 1621 CTTGACGGTGGCGGCCCTCTCTCCACCATCCAGAGTACTCGAAAGACCTTCTTTGTCCTGCGG 1680
Db 1666 CTTGACGGTGGCGGCCCTCTCTCCACCATCCAGAGTACTCGAAAGACCTTCTTTGTCCTGCGG 1725
QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATGCCCGGCAACCGGTCGCT 1740
Db 1726 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATGCCCGGCAACCGGTCGCT 1785
QY 1741 ATTTCACCTTACACCACTAGCCTGGGTGCTGGTCCCGTCTCCATTTCTCGGTTGCCGTT 1800
Db 1786 ATTTCACCTTACACCACTAGCCTGGGTGCTGGTCCCGTCTCCATTTCTCGGTTGCCGTT 1845
QY 1801 TTAGCCCCCACTCTCGGCTAGCATTTGCTTGGAGTACCTTGGACTACCTGCGCGGCC 1860
Db 1846 TTAGCCCCCACTCTCGGCTAGCATTTGCTTGGAGTACCTTGGACTACCTGCGCGGCC 1905
QY 1861 CATACTTTTGTGATTTCTGCGGAGTCCGCGGCTTGGCTTGGCTTCCAGGCTGCGCTTTC 1920
Db 1906 CATACTTTTGTGATTTCTGCGGAGTCCGCGGCTTGGCTTGGCTTCCAGGCTGCGCTTTC 1965
QY 1921 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTAAACTCGGGAGTTG 1980
Db 1966 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTAAACTCGGGAGTTG 2025
QY 1981 TAG 1983
Db 2026 TAG 2028

RESULT 12

US-07-870-985A-1
; Sequence 1, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Two, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; Zip: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888

```
/ FILING DATE: 05-APRIL-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 420,921
/ FILING DATE: 13-OCTOBER-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 367,486
/ FILING DATE: 16-JUNE-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 336,672
/ FILING DATE: 11-APRIL-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 208,997
/ FILING DATE: 17-JUNE-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0093.30
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2094 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA to mRNA
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: BURMA SEQUENCE, FIGURE 7
/ US-07-870-985A-1

Query Match 100.0%; Score 1983; DB 4; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCGGCGCTATTTGTGCTGCTCTCATGCTTTTGGCCATGCTGCCCGCGCA 60
DB 46 ATGCGCCCTCGGCGCTATTTGTGCTGCTCTCATGCTTTTGGCCATGCTGCCCGCGCA 105
QY 61 CGCGCCGCTGAGCGCTCTGCGCGCGCTGCGCGCGCGCGCGCGCGCTGCGCGCGT 120
DB 106 CGCGCCGCTGAGCGCTCTGCGCGCGCTGCGCGCGCGCGCGCGCGCTGCGCGCGT 165
QY 121 TTCTGGGCTGACCGGTTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC 180
DB 166 TTCTGGGCTGACCGGTTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC 225
QY 181 CCCTTCGCGCCCGGATGTACCGCTCGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 240
DB 226 CCCTTCGCGCCCGGATGTACCGCTCGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 285
QY 241 CGACACTCGGCTCGCTTGGGTGACCGAGCCCAAGCGCCCGCGCTTGCCTTCAGTCT 300
DB 286 CGACACTCGGCTCGCTTGGGTGACCGAGCCCAAGCGCCCGCGCTTGCCTTCAGTCT 345
QY 301 AGACCTACACAGCTGGCGCGCGCTTACCGCGGTGCTCGCGCGCCATGACACCCGG 360
DB 346 AGACCTACACAGCTGGCGCGCGCTTACCGCGGTGCTCGCGCGCCATGACACCCGG 405
QY 361 CCAGTGCTGATGTGCGACTCCCGCGCGCCATTTTCGCCCGGCGAGTAACTATCAACA 420
DB 406 CCAGTGCTGATGTGCGACTCCCGCGCGCCATTTTCGCCCGGCGAGTAACTATCAACA 465
QY 421 TCTCCCTTACCTTTCGCGCGCCACCGGCACTAACCTGGTCTTTATGCCGCCCTCTT 480
DB 466 TCTCCCTTACCTTTCGCGCGCCACCGGCACTAACCTGGTCTTTATGCCGCCCTCTT 525
QY 481 AGTCGGCTTTTACCCCTTCAGGACGGCAACCAATACCATATAATGGCCACGAGCTTCT 540
DB 526 AGTCGGCTTTTACCCCTTCAGGACGGCAACCAATACCATATAATGGCCACGAGCTTCT 585
QY 541 AATTATGCCAGTACCGGGTTGCCCGTGCGCAATCCCGTTACCGCGCGCTGGTCCCAAT 600
```

```
DB 586 AATTATGCCAGTACCGGGTTGCCCGTGCGCAATCCGTACCGCGCGTGGTCCCAAT 645
QY 601 GCTGTGCGGGTTACGCCATCTCCATCTCATTTCTGGCCACAGACCAACCCCGGCG 660
DB 646 GCTGTGCGGGTTACGCCATCTCCATCTCATTTCTGGCCACAGACCAACCCCGGCG 705
QY 661 TCCGTTGATATGAATTCATTAACCTCGAGGAGTTCGTATTTTAGTCAGCGCGGCATA 720
DB 706 TCCGTTGATATGAATTCATTAACCTCGAGGAGTTCGTATTTTAGTCAGCGCGGCATA 765
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTTACACTATCGTAAACCAAGGCTGGCGCTCC 780
DB 766 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTTACACTATCGTAAACCAAGGCTGGCGCTCC 825
QY 781 GTCGAGACTCTTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTTGTTATGCTTTGCATA 840
DB 826 GTCGAGACTCTTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTTGTTATGCTTTGCATA 885
QY 841 CATGGCTCACTCGTAAATTCCTATATACTAATAACCCCTATACCGGTGCCCTCGGGCTGTG 900
DB 886 CATGGCTCACTCGTAAATTCCTATATACTAATAACCCCTATACCGGTGCCCTCGGGCTGTG 945
QY 901 GACTTTGGCCCTTGAGCTTGAAGTTTCGCAACCTTACCCCGGTAACACCAATACGCGGCTC 960
DB 946 GACTTTGGCCCTTGAGCTTGAAGTTTCGCAACCTTACCCCGGTAACACCAATACGCGGCTC 1005
QY 961 TCCCGTTATTCAGCACTGTCTGCGCACCGCTTGTGCGGGTGGGAGCGGAGTCCCGAG 1020
DB 1006 TCCCGTTATTCAGCACTGTCTGCGCACCGCTTGTGCGGGTGGGAGCGGAGTCCCGAG 1065
QY 1021 CTCACCAACGCGCTGCTACCGCTTTATGAGGAGCTCTATTTTACTAGTACTAATGTT 1080
DB 1066 CTCACCAACGCGCTGCTACCGCTTTATGAGGAGCTCTATTTTACTAGTACTAATGTT 1125
QY 1081 GTCGCTGAGATCGGCGCGGATAGCCCTCACTCCCTGTTTCAACCTTGTGACACTCTGCTT 1140
DB 1126 GTCGCTGAGATCGGCGCGGATAGCCCTCACTCCCTGTTTCAACCTTGTGACACTCTGCTT 1185
QY 1141 GCGGCGCTGCGACAGAAATGATTTGTCGCGCTGGTGCGGAGCTGTTTACTTCCCGTCCC 1200
DB 1186 GCGGCGCTGCGACAGAAATGATTTGTCGCGCTGGTGCGGAGCTGTTTACTTCCCGTCCC 1245
QY 1201 GTTGTCTCAGCAATGCGCGCGAGCTGTTAAGTTGTATACATCTGTAGAGAGTCTCAG 1260
DB 1246 GTTGTCTCAGCAATGCGCGAGCTGTTAAGTTGTATACATCTGTAGAGAGTCTCAG 1305
QY 1261 CAGGATAAGGCTATTTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGTTATT 1320
DB 1306 CAGGATAAGGCTATTTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGTTATT 1365
QY 1321 CAGGATAAGTAAACCAACATGAACATGCGGCGAGCTTCTTCCAGCCCGCATTCGCGC 1380
DB 1366 CAGGATAAGTAAACCAACATGAACATGCGGCGAGCTTCTTCCAGCCCGCATTCGCGC 1425
QY 1381 CCTTCTCTGCTTCGAGCTAATGATGCTGCTTTGGCTCTCTCAGCGCTCGGAGTAT 1440
DB 1426 CCTTCTCTGCTTCGAGCTAATGATGCTGCTTTGGCTCTCTCAGCGCTCGGAGTAT 1485
QY 1441 GACCAGTCCACTTATGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500
DB 1486 GACCAGTCCACTTATGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1545
QY 1501 GTTAAATGTCGACGCGCGCGAGCGGCTTGCCTGGCTGCTGATTTGGAACAAGGTCA 1560
DB 1546 GTTAAATGTCGACGCGCGCGAGCGGCTTGCCTGGCTGCTGATTTGGAACAAGGTCA 1605
QY 1561 CTTGACCGTTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCTCGCG 1620
DB 1606 CTTGACCGTTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCTCGCG 1665
QY 1621 CTTCCGCGTAACTCTCTTTCTGGGAGCGGCAACCAAGCCGGTACCTTATATAT 1680
```

Db 1666 CTCGCGGTAGCTCTCTTCTGGGAGGAGGACAACTAAAGCCGGTACCTTATAAT 1725
QY 1681 TATAACACACTGCTAGGACCAACTGCTTGTGAGAAATGCCCGGACCGGGTCGCT 1740
Db 1726 TATAACACACTGCTAGGACCAACTGCTTGTGAGAAATGCCCGGACCGGGTCGCT 1785
QY 1741 ATTTCACATTACACCACTAGCTGGTGGTCCCGTCTCCATTTCTGCGGTGCGTT 1800
Db 1786 ATTTCACATTACACCACTAGCTGGTGGTCCCGTCTCCATTTCTGCGGTGCGTT 1845
QY 1801 TTAGCCCCCACTCTGCGTAGCAATGCTTGAGGATACCTTGGAATACCTGCGCGCC 1860
Db 1846 TTAGCCCCCACTCTGCGTAGCAATGCTTGAGGATACCTTGGAATACCTGCGCGCC 1905
QY 1861 CATACTTTTGATGATTTCTGCCAGAGTGCAGGCTGCGGCTTTCAGGGCTGCGCTTC 1920
Db 1906 CATACTTTTGATGATTTCTGCCAGAGTGCAGGCTGCGGCTTTCAGGGCTGCGCTTC 1965
QY 1921 CAGTCTACTGCTGCTGAGCTTACAGCCCTTAAGATGAAGTGGGTAAACTCGGAGTTG 1980
Db 1965 CAGTCTACTGCTGCTGAGCTTACAGCCCTTAAGATGAAGTGGGTAAACTCGGAGTTG 2025
QY 1981 TAG 1983
Db 2026 TAG 2028

RESULT 13

US-08-478-507-6
; Sequence 6, Application US/08478507
; Patent No. 6120988
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarborough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6120988-A/No. 6120988-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,507
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672

; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0183.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HEV - Burma strain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..5106
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5147..7126
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5106..5474
; US-08-478-507-6

Query Match 100.0%; Score 1983; DB 3; Length 7195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCGGCTATTTTGTGCTCCTCATGTTTGTGCTATGCTGCCGCGGCCA 60
Db 5147 ATGCGCCCTCGGCTATTTTGTGCTCCTCATGTTTGTGCTATGCTGCCGCGGCCA 5206

QY 61 CCGCCCGGTGAGCCGCTCGTGGCGCGCGAGCGCGGTTCGGCGGTGCT 120
Db 5207 CCGCCCGGTGAGCCGCTCGTGGCGCGCGAGCGCGGTTCGGCGGTGCT 5266

QY 121 TTCTGGGTGAGCGGGTTCATTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 5267 TTCTGGGTGAGCGGGTTCATTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 5326

QY 181 CCTTCGCCCGCGATGTCAACCGCTGCGCGCGGGGCTGGACCTGCTGTTCCGCCAACCCGCC 240
Db 5327 CCTTCGCCCGCGATGTCAACCGCTGCGCGCGGGGCTGGACCTGCTGTTCCGCCAACCCGCC 5386

QY 241 CGACCACTCGGTTCGCTTGGCGGTGACAGGCGCCAGGCGCGCGGTTCGCTCAGCTCGT 300
Db 5387 CGACCACTCGGTTCGCTTGGCGGTGACAGGCGCCAGGCGCGCGGTTCGCTCAGCTCGT 5446

QY 301 AGACCTACACAGCTGCGGCGCGCGCTAACCGCGGTGCTCCGCGCCCATGACACCCCG 360
Db 5447 AGACCTACACAGCTGCGGCGCGCGCTAACCGCGGTGCTCCGCGCCCATGACACCCCG 5506

QY 361 CCAGTGCCTGATGTGACTCCCGCGCGCGCATCTTCGCGCGGCGAGTATACCTTATCAACA 420
Db 5507 CCAGTGCCTGATGTGACTCCCGCGCGCGCATCTTCGCGCGGCGAGTATACCTTATCAACA 5566

QY 421 TCTCCCTTACCTCTTCCGTGGCGACCGGCACTAACTGGTTCTTTTATGCGCCCTCTT 480
Db 5567 TCTCCCTTACCTCTTCCGTGGCGACCGGCACTAACTGGTTCTTTTATGCGCCCTCTT 5626

QY 481 AGTCCGCTTTTACCCCTTCAGGACGCGCACCAATACCATATATATGCGCACGAGCTTCT 540
Db 5627 AGTCCGCTTTTACCCCTTCAGGACGCGCACCAATACCATATATATGCGCACGAGCTTCT 5686

QY 541 AATTATGCCAGTACCGGGTGGCCGTCGACCAATCCGTTACCGCCGCGTGGTCCCAAT 600
DB 5687 AATTATGCCAGTACCGGGTGGCCGTCGACCAATCCGTTACCGCCGCGTGGTCCCAAT 5746
QY 601 GCTGTCGGGGTACCGCACTCCATCTCATCTCTGCGCCACAGACCAACACCCCGAAG 660
DB 5747 GCTGTCGGGGTACCGCACTCCATCTCATCTCTGCGCCACAGACCAACACCCCGAAG 5806
QY 661 TCCGTCGATATGAATCAATTAACCTCGACGAGTCTGTTATTTAGTCCAGCCCGGCATA 720
DB 5807 TCCGTCGATATGAATCAATTAACCTCGACGAGTCTGTTATTTAGTCCAGCCCGGCATA 5866
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGTGCGGTCC 780
DB 5867 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGTGCGGTCC 5926
QY 781 GTCGAGACCTTGGGGTGGTGAGGAGAGGCTACCTCTGCTGCTTGTATGCTTTGCA 840
DB 5927 GTCGAGACCTTGGGGTGGTGAGGAGAGGCTACCTCTGCTGCTTGTATGCTTTGCA 5986
QY 841 CATGGCTCACTCGTAATTCCTATCTATACTAATAACCCCTATACCGGTGCCCTCGGGCTGTG 900
DB 5987 CATGGCTCACTCGTAATTCCTATCTATACTAATAACCCCTATACCGGTGCCCTCGGGCTGTG 6046
QY 901 GACTTTGCCCTTGAGCTTGATTTGCAACCTTACCCCGGTAAACCAATACCGGGTC 960
DB 6047 GACTTTGCCCTTGAGCTTGATTTGCAACCTTACCCCGGTAAACCAATACCGGGTC 6106
QY 961 TCCGTTATTCAGACTGCTGCGCACCGCTTCTGTCGGTGGGACGGAGCTGCGAG 1020
DB 6107 TCCGTTATTCAGACTGCTGCGCACCGCTTCTGTCGGTGGGACGGAGCTGCGAG 6166
QY 1021 CTCACCAACGCGTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAAGT 1080
DB 6167 CTCACCAACGCGTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAAGT 6286
QY 1081 GTCGGTGAGATCGCCGCGGATAGCCCTCAACCTGTTTCAACCTTGCTGACACTGCTT 1140
DB 6227 GTCGGTGAGATCGCCGCGGATAGCCCTCAACCTGTTTCAACCTTGCTGACACTGCTT 6286
QY 1141 GCGGCTCGCGACAGAAATGATTTGCTGCGTGGTGGCCAGCTGTTCTACTCCCGTCC 1200
DB 6287 GCGGCTCGCGACAGAAATGATTTGCTGCGTGGTGGCCAGCTGTTCTACTCCCGTCC 6346
QY 1201 GTTCTCTCAGCAATGCGAGCCAGTGAAGTTGTATACATCTGTAGAGAAGTCTCAG 1260
DB 6347 GTTCTCTCAGCAATGCGAGCCAGTGAAGTTGTATACATCTGTAGAGAAGTCTCAG 6406
QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT 1320
DB 6407 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT 6466
QY 1321 CAGGATATGATACCAACATGAACAGATCGCGGAGCGCTTCTCAGCCCGCATCGCGC 1380
DB 6467 CAGGATATGATACCAACATGAACAGATCGCGGAGCGCTTCTCAGCCCGCATCGCGC 6526
QY 1381 CTTTCTCTGCTCTGAGCTTAATGATGCTGTTGGCTCTCTCTCAGCGTGGCGAGTAT 1440
DB 6527 CTTTCTCTGCTCTGAGCTTAATGATGCTGTTGGCTCTCTCTCAGCGTGGCGAGTAT 6586
QY 1441 GACCAGTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGATCTGTGACCTTG 1500
DB 6587 GACCAGTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGATCTGTGACCTTG 6646
QY 1501 GTTAATGTTGGACCGGGCGAGCGCTTCCCGGTGCTGATTTGAGTGAACCAAGTCA 1560
DB 6647 GTTAATGTTGGACCGGGCGAGCGCTTCCCGGTGCTGATTTGAGTGAACCAAGTCA 6706
QY 1561 CTTGACCGTCCCGCTCTCCACCATCCAGAGTACTCGAGAGCTTCTTTGCTCGCG 1620
DB 6707 CTTGACCGTCCCGCTCTCCACCATCCAGAGTACTCGAGAGCTTCTTTGCTCGCG 6766
QY 1621 CTCGCGGTAAGCTCTCTTTCTGGGAGGAGGCAACTAAAGCGGGTACCTTTATAAT 1680

DB 6767 CTCGCGGTAAGCTCTCTTTCTGGAGGAGGCAAACTAAAGCGGGTACCTTTATAAT 6826
QY 1681 TATAACACCACTGCTAGGACCAACTGCTTGTGAGAAATGCGCGGACCGGGTGGCT 1740
DB 6827 TATAACACCACTGCTAGGACCAACTGCTTGTGAGAAATGCGCGGACCGGGTGGCT 6886
QY 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTGGGGTGGCGTT 1800
DB 6887 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTGGGGTGGCGTT 6946
QY 1801 TTAGCCCCCACTCTGCGCTAGCACTTCTTGAAGATACCTTGGACTACCTTCCCGCGCC 1860
DB 6947 TTAGCCCCCACTCTGCGCTAGCACTTCTTGAAGATACCTTGGACTACCTTCCCGCGCC 7006
QY 1861 CATACTTTTATGATTTCTGCCAGAGTGGCGCCCTTGGCCCTTCAGGGTGGCGTTTTC 1920
DB 7007 CATACTTTTATGATTTCTGCCAGAGTGGCGCCCTTGGCCCTTCAGGGTGGCGTTTTC 7066
QY 1921 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTAATACTCGGGAGTTG 1980
DB 7067 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTAATACTCGGGAGTTG 7126
QY 1981 TAG 1983
DB 7127 TAG 7129

RESULT 14
US-09-128-275A-6
; Sequence 6, Application US/09128275A
; Patent No. 6229005
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6229005-A/No. 6229005-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,275A
; FILING DATE: 03-AUG-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HEV - Burma strain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..5106
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5147..7126
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5106..5474
; US-09-128-275A-6

Query Match 100.0%; Score 1983; DB 3; Length 7195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCGCCCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCGCGGCA 60
DB 5147 ATGCGCCCTCGCCCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCGCGGCA 5206

QY 61 CGCGCCGCTCAGCGCTGCGCGCGCGCTGCTGCGCGCGCGCAGCGCGGTTCCGCGCGTGT 120
DB 5207 CGCGCCGCTCAGCGCTGCGCGCGCGCTGCTGCGCGCGCGCAGCGCGGTTCCGCGCGTGT 5266

QY 121 TTTCTGGGTGACCGGGTTGATTTCTCAGCCCTTTCGCAATCCCTATATTCATCAACCAAC 180
DB 5267 TTTCTGGGTGACCGGGTTGATTTCTCAGCCCTTTCGCAATCCCTATATTCATCAACCAAC 5326

QY 181 CCCTTCGCCCGCGATGTACCGCTGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 240
DB 5327 CCCTTCGCCCGCGATGTACCGCTGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 5386

QY 241 CGACCACTCGGCTCGGCTGGCGTGACCAAGGCCCGCGCGCGCGCTTGCCTCAGCTCGT 300
DB 5387 CGAACACTCGGCTCGGCTGGCGTGACCAAGGCCCGCGCGCGCGCTTGCCTCAGCTCGT 5446

QY 301 AGACCTACACAGCTGGGCGCGCGCTAAACCGCGTGTCTCGCGCGCGCGCGCGCG 360
DB 5447 AGACCTACACAGCTGGGCGCGCGCTAAACCGCGTGTCTCGCGCGCGCGCGCGCG 5506

QY 361 CCAAGTGCCTGATGTGACCTCCCGCGCGCATCTTTCGCGCGCGCATATAACCTATCAACA 420
DB 5507 CCAAGTGCCTGATGTGACCTCCCGCGCGCATCTTTCGCGCGCGCATATAACCTATCAACA 5566

QY 421 TCTCCCTTACCTTTCGCGTGGCGCGCGCGCATAAACCTGGTTCCTTATGCGCGCGCTCT 480
DB 5567 TCTCCCTTACCTTTCGCGTGGCGCGCGCGCATAAACCTGGTTCCTTATGCGCGCGCTCT 5626

QY 481 AGTCCGCTTTTACCCCTTCAGACCGCGCACCAATACCATATATAATGGCGCACGAGCTTCT 540
DB 5627 AGTCCGCTTTTACCCCTTCAGACCGCGCACCAATACCATATATAATGGCGCACGAGCTTCT 5686
```


QY 481 AGTCCGCTTTTACCCCTTCAGGACGGCACAATACCATATATATATGCGCACGGAAGCTTCT 540
 Db
 5627 AGTCCGCTTTTACCCCTTCAGGACGGCACAATACCATATATATATGCGCACGGAAGCTTCT 5686
 QY 541 AATTATGCCAGTACCGGGTTCGCCGTGCACAAATCCGTTACCGCCCGCTGTGTCGCCAAT 600
 Db 5687 AATTATGCCAGTACCGGGTTCGCCGTGCACAAATCCGTTACCGCCCGCTGTGTCGCCAAT 5746
 QY 601 GCTGTGCGGGTTACCGCATCTCCATCTCATTTCTGGCCACAGACCAACCAACCCCGACG 660
 Db 5747 GCTGTGCGGGTTACCGCATCTCCATCTCATTTCTGGCCACAGACCAACCAACCCCGACG 5806
 QY 661 TCCGTTGATATGAATTAATTAACCTCGACGAGTGTTCGTATTTAGTCCAGCCCGGCATA 720
 Db 5807 TCCGTTGATATGAATTAATTAACCTCGACGAGTGTTCGTATTTAGTCCAGCCCGGCATA 5866
 QY 721 GCTCTGAGCTTGATGCCAAGTGAAGCGCTACACTATCGTAAACAAAGCTGGCGCTCC 780
 Db 5867 GCTCTGAGCTTGATGCCAAGTGAAGCGCTACACTATCGTAAACAAAGCTGGCGCTCC 5926
 QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGCTCTTCTGTTATGCTTGCATA 840
 Db 5927 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGCTCTTCTGTTATGCTTGCATA 5986
 QY 841 CATGGCTCACTCGTAAATTCCTATATACTAATACACCTTATACCGGTGCCCTCGGGCTGTTG 900
 Db 5987 CATGGCTCACTCGTAAATTCCTATATACTAATACACCTTATACCGGTGCCCTCGGGCTGTTG 6046
 QY 901 GACTTTGCCCTTGAGCTTGAGTTTGGCAACTTACCCCGGTAAACCAATACCGGGTC 960
 Db 6047 GACTTTGCCCTTGAGCTTGAGTTTGGCAACTTACCCCGGTAAACCAATACCGGGTC 6106
 QY 961 TCCCGTTATTCAGCACTGCTCCGACCGCTTCGTCGGGTGCGAGCGGACTGCGGAG 1020
 Db 6107 TCCCGTTATTCAGCACTGCTCCGACCGCTTCGTCGGGTGCGAGCGGACTGCGGAG 6166
 QY 1021 CTCACCAACACCGCTGTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080
 Db 6167 CTCACCAACACCGCTGTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 6226
 QY 1081 GTCGGTGAGATCGGCGCGGGATAGCCCTCACCTGTTCAACCTTGTGACACCTCTGCTT 1140
 Db 6227 GTCGGTGAGATCGGCGCGGGATAGCCCTCACCTGTTCAACCTTGTGACACCTCTGCTT 6286
 QY 1141 GCGGGCTGCGGACAGAAATTTGTTTCTGTCGGCTGTGGCCAGCTGTTCTACTCCCGTCCC 1200
 Db 6287 GCGGGCTGCGGACAGAAATTTGTTTCTGTCGGCTGTGGCCAGCTGTTCTACTCCCGTCCC 6346
 QY 1201 GTTGTCTCAGCCAATGGCGCGGACTGTTTAACTTGTATATACATCTGTAGAGAATGCTCAG 1260
 Db 6347 GTTGTCTCAGCCAATGGCGCGGACTGTTTAACTTGTATATACATCTGTAGAGAATGCTCAG 6406
 QY 1261 CAGGATAAGGGTATTGGAATCCCGCATGACATTTGACCTCGGAGAACTCTGTTGGTTATT 1320
 Db 6407 CAGGATAAGGGTATTGGAATCCCGCATGACATTTGACCTCGGAGAACTCTGTTGGTTATT 6466
 QY 1321 CAGGATTATGATAACCAACATGAACAGATCGCGCGAGCGCTTCTCCAGCCCATCGCGC 1380
 Db 6467 CAGGATTATGATAACCAACATGAACAGATCGCGCGAGCGCTTCTCCAGCCCATCGCGC 6526
 QY 1381 CCTTTCTCTCTCTCGAGCTAATGATGTSCTTTGGCTCTCTCTCACCCTGCGCGAGTAT 1440
 Db 6527 CCTTTCTCTCTCTCGAGCTAATGATGTSCTTTGGCTCTCTCTCACCCTGCGCGAGTAT 6586
 QY 1441 GACCACTCAGCTTATGGCTCTTCCAGTGGCGCGAGTTATGTTTCTGACTCTGACCTTG 1500
 Db 6587 GACCACTCAGCTTATGGCTCTTCCAGTGGCGCGAGTTATGTTTCTGACTCTGACCTTG 6646
 QY 1501 GTTAATGTTGCGACCGCGCGAGCGGCTTCCCGGTGCGCTCGATTGGACCAAGGTCACA 1560
 Db 6647 GTTAATGTTGCGACCGCGCGAGCGGCTTCCCGGTGCGCTCGATTGGACCAAGGTCACA 6706
 QY 1561 CTTGACGGTGGCCCTCTCCACCAATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGCG 1620

Db 6707 CTTGACGGTGGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCCTGCGG 6766
 QY 1621 CTCGCGGTTAAGCTCTCTTTCTGGGAGGAGGACGACAACTAAAGCCGGGTACCCCTTATAAT 1680
 Db 6767 CTCGCGGTTAAGCTCTCTTTCTGGGAGGAGGACGACAACTAAAGCCGGGTACCCCTTATAAT 6826
 QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATGCGCGCGGACACCGGGTCGCT 1740
 Db 6827 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATGCGCGCGGACACCGGGTCGCT 6886
 QY 1741 AATTCCACTTTACACCACTAGCCCTGGGTGCTGCTGCCGTCTCCATTTCTGCGGTTGCCGTT 1800
 Db 6887 AATTCCACTTTACACCACTAGCCCTGGGTGCTGCTGCCGTCTCCATTTCTGCGGTTGCCGTT 6946
 QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGAGACTACCCCTGCCCGGCC 1860
 Db 6947 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGAGACTACCCCTGCCCGGCC 7006
 QY 1861 CATACTTTTGTGATTTCTGCCAGAGTGGCGCCCTTGGCCTTCAGGGCTGCGCTTTC 1920
 Db 7007 CATACTTTTGTGATTTCTGCCAGAGTGGCGCCCTTGGCCTTCAGGGCTGCGCTTTC 7066
 QY 1921 CAGTCTACTGTCGCTGAGCTTCAGCGCTTAAAGATGAAGGTGGGTAAACTCGGGAGTTG 1980
 Db 7067 CAGTCTACTGTCGCTGAGCTTCAGCGCTTAAAGATGAAGGTGGGTAAACTCGGGAGTTG 7126
 QY 1981 TAG 1983
 Db 7127 TAG 7129

Search completed: August 21, 2004, 18:10:41
 Job time : 159.606 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 15:15:13 ; Search time 886.671 Seconds
(without alignments)
10985.769 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129

Perfect score: 1983
Sequence: 1 ATGCGCCTCGGCTATT...GTAAACTCGGAGTGTAG 1983

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3228839 seqs, 245606551 residues

Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1983	100.0	2049	9	US-09-769-066-1
2	1983	100.0	2049	15	US-10-165-868-3
3	1983	100.0	2094	15	US-10-165-868-1
4	1983	100.0	7195	10	US-09-851-410-6
5	1805.4	91.0	1990	13	US-10-381-770-5
6	1793.4	90.4	1989	13	US-10-381-770-6
7	1647	83.1	1647	9	US-09-769-066-3
8	1384.6	69.8	2058	9	US-09-769-066-2
9	1368.6	69.0	2055	15	US-10-165-868-4
10	1368.6	69.0	2100	15	US-10-165-868-2
11	1368.6	69.0	7171	10	US-09-851-410-10
12	1368.6	69.0	7180	17	US-10-239-090A-49
13	1367.6	69.0	7230	17	US-10-239-090A-48
14	1347.6	68.0	7228	17	US-10-239-090A-47

Sequence 45, Appl
Sequence 11, Appl
Sequence 46, Appl
Sequence 44, Appl
Sequence 164, App
Sequence 165, App
Sequence 164, App
Sequence 89, Appl
Sequence 165, App
Sequence 90, Appl
Sequence 89, Appl
Sequence 90, Appl
Sequence 9, Appl
Sequence 4, Appl
Sequence 5, Appl
Sequence 7, Appl
Sequence 10, Appl
Sequence 129, App
Sequence 129, App
Sequence 6, Appl
Sequence 8, Appl
Sequence 193, App
Sequence 196, App
Sequence 193, App
Sequence 196, App
Sequence 192, App
Sequence 195, App
Sequence 192, App
Sequence 195, App
Sequence 24, Appl
Sequence 24, Appl

15 1346.8 67.9 7233 17 US-10-239-090A-45
16 1346 67.9 7234 17 US-10-239-090A-11
17 1344.4 67.8 7239 17 US-10-239-090A-46
18 1343.6 67.8 7256 17 US-10-239-090A-44
19 1339.6 67.6 7277 10 US-09-468-147-164
20 1339.6 67.6 7277 10 US-09-468-147-165
21 1339.6 67.6 7277 13 US-10-319-745-164
22 1339.6 67.6 7277 13 US-10-319-745-165
23 1323.2 66.7 7202 10 US-09-468-147-89
24 1323.2 66.7 7202 10 US-09-468-147-90
25 1323.2 66.7 7202 13 US-10-319-745-89
26 1323.2 66.7 7202 13 US-10-319-745-90
27 1311 66.1 1311 15 US-10-165-868-9
28 1106.2 55.8 1647 9 US-09-769-066-4
29 984 49.6 984 9 US-09-769-066-5
30 984 49.6 984 15 US-10-165-868-7
31 880.6 44.4 1308 15 US-10-165-868-10
32 675.6 34.1 1020 10 US-09-468-147-129
33 675.6 34.1 1020 13 US-10-319-745-129
34 673.6 34.0 984 9 US-09-769-066-6
35 672.2 33.9 981 15 US-10-165-868-8
36 658.8 33.2 1389 10 US-09-468-147-193
37 658.8 33.2 1389 10 US-09-468-147-196
38 658.8 33.2 1389 13 US-10-319-745-193
39 658.8 33.2 1389 13 US-10-319-745-196
40 652.6 32.9 1026 10 US-09-468-147-192
41 652.6 32.9 1026 10 US-09-468-147-195
42 652.6 32.9 1026 13 US-10-319-745-192
43 652.6 32.9 1026 13 US-10-319-745-195
44 583.8 29.4 899 10 US-09-468-147-24
45 583.8 29.4 899 13 US-10-319-745-24

ALIGNMENTS

RESULT 1

US-09-769-066-1
; Sequence 1, Application US/09769066
; Patent No. US2020107360A1

GENERAL INFORMATION:

APPLICANT: Fuerst, Thomas R.
McAttee, C. Patrick
Yarborough, Patrice O.
Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Ave., Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/769,066

FILING DATE: 24-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/542,634

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 4600-0293.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

```
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2049 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: Hepatitis E Virus (Burma strain)
;     TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;     HYPOTHETICAL: NO
;     ORIGINAL SOURCE:
;     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-769-066-1

Query Match      100.0%; Score 1983; DB 9; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGCGGCTCGGCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCGCGCCA 60
DB      1 ATGCGGCTCGGCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCGCGCCA 60

QY      61 CCGCCCGGTCAGCGCTGTCGCGCGCTGCGGCGCGGCGAGCGCGGTCGCGGGTGGT 120
DB      61 CCGCCCGGTCAGCGCTGTCGCGCGCTGCGGCGCGGCGAGCGCGGTCGCGGGTGGT 120

QY      121 TTCTGCGGTTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTCATATTTATCCCAAC 180
DB      121 TTCTGCGGTTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTCATATTTATCCCAAC 180

QY      181 CCTTCGCCCCGATGTCACCGCTGCGCGCGGCTGAGCTGCTGTCGCCAACCCGCG 240
DB      181 CCTTCGCCCCGATGTCACCGCTGCGCGCGGCTGAGCTGCTGTCGCCAACCCGCG 240

QY      241 CGAACCTCTCGGCTCCGCTTGGCGTGACAGGCCAGCGCCGCGCTGCGGCCCATGACACCCG 300
DB      241 CGAACCTCTCGGCTCCGCTTGGCGTGACAGGCCAGCGCCGCGCTGCGGCCCATGACACCCG 300

QY      301 AGACTTACACAGCTGGGCGCGCGCTAACCGCGGTGCGTCCGCGCCCATGACACCCG 360
DB      301 AGACTTACACAGCTGGGCGCGCGCTAACCGCGGTGCGTCCGCGCCCATGACACCCG 360

QY      361 CCAGTGCCTGATGTCAGCTCCGCGCGCGCATCTTGGCGCGGAGTAACTATCAACA 420
DB      361 CCAGTGCCTGATGTCAGCTCCGCGCGCGCATCTTGGCGCGGAGTAACTATCAACA 420

QY      421 TCTCCGCTTACCTCTTCCGTGCGCACCGGCACTAACCTGCTTCTTATGCGCGCCCTCT 480
DB      421 TCTCCGCTTACCTCTTCCGTGCGCACCGGCACTAACCTGCTTCTTATGCGCGCCCTCT 480

QY      481 AGTCGCTTTTACCGCTTCAGGACGGCACCAATACCCATATATATGCGCACGGAAGCTTCT 540
DB      481 AGTCGCTTTTACCGCTTCAGGACGGCACCAATACCCATATATATGCGCACGGAAGCTTCT 540

QY      541 AATATGCCAGTACCGGTTGCGCGTGCCGACCAATCCGTTACCGCGCGCTGTCGCCAAT 600
DB      541 AATATGCCAGTACCGGTTGCGCGTGCCGACCAATCCGTTACCGCGCGCTGTCGCCAAT 600

QY      601 GCTGTCGCGGTTAGCCATCTCCATCTCACTTCTGCGCACAGACCAACACCCCGGAG 660
DB      601 GCTGTCGCGGTTAGCCATCTCCATCTCACTTCTGCGCACAGACCAACACCCCGGAG 660

QY      661 TCCGTTGATATGAATTAATTAACCTCGAGCGGATGTTGTTATGTCAGCCGCGCATA 720
DB      661 TCCGTTGATATGAATTAATTAACCTCGAGCGGATGTTGTTATGTCAGCCGCGCATA 720

QY      721 GCCTCTGAGCTTGATCCCAAGTACGCGCTTACACTATCGTAAACCAAGGTCGGCTCC 780
DB      721 GCCTCTGAGCTTGATCCCAAGTACGCGCTTACACTATCGTAAACCAAGGTCGGCTCC 780

QY      781 GTCGAGACCTCTGGGCTGGCTGAGGAGGAGGCTACCTCTGCTCTTGTATGCTTGCATA 840
DB      781 GTCGAGACCTCTGGGCTGGCTGAGGAGGAGGCTACCTCTGCTCTTGTATGCTTGCATA 840
```

```
QY      841 CATGGCTCACTCGTAATTCCTATACATAACACCTATACCGGTGCGCTCGGGCTGTG 900
DB      841 CATGGCTCACTCGTAATTCCTATACATAACACCTATACCGGTGCGCTCGGGCTGTG 900

QY      901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGTC 960
DB      901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGTC 960

QY      961 TCCGTTATTCAGCACTGCTGCCACCGCTTTCGTGCGGTGCGGACGCGACTGCCGAG 1020
DB      961 TCCGTTATTCAGCACTGCTGCCACCGCTTTCGTGCGGTGCGGACGCGACTGCCGAG 1020

QY      1021 CTCACCAACGAGCTGCTACCGCTTATGAAAGACCTCTATTCTAGTACTATGCT 1080
DB      1021 CTCACCAACGAGCTGCTACCGCTTATGAAAGACCTCTATTCTAGTACTATGCT 1080

QY      1081 GTCGCTGAGATCGCGCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
DB      1081 GTCGCTGAGATCGCGCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140

QY      1141 GCGCGCTCGCGACAGATTGATTTCGTCGGCTGCTGGCGAGCTGTTCTACTCCGTC 1200
DB      1141 GCGCGCTCGCGACAGATTGATTTCGTCGGCTGCTGGCGAGCTGTTCTACTCCGTC 1200

QY      1201 GTTGTCTCAGCCAAATGCGGAGCGGACTGTTAAGTTGTTATACATCTCTAGAGATGCTCAG 1260
DB      1201 GTTGTCTCAGCCAAATGCGGAGCGGACTGTTAAGTTGTTATACATCTCTAGAGATGCTCAG 1260

QY      1261 CAGGATAAGGGTATTGCAATCCGATGACATTTGACCTCGGAGAACTCTGCTGCTGTTATT 1320
DB      1261 CAGGATAAGGGTATTGCAATCCGATGACATTTGACCTCGGAGAACTCTGCTGCTGTTATT 1320

QY      1321 CAGGATTATGATAACCAATGAAACAGATCGCGCGAGCGCTTCTCCAGCCCATCGCG 1380
DB      1321 CAGGATTATGATAACCAATGAAACAGATCGCGCGAGCGCTTCTCCAGCCCATCGCG 1380

QY      1381 CTTTCTCTGCTTTCGAGCTAAATGATGCTGCTTGGCTCTCTCTACCGCTGCGAGTAT 1440
DB      1381 CTTTCTCTGCTTTCGAGCTAAATGATGCTGCTTGGCTCTCTCTACCGCTGCGAGTAT 1440

QY      1441 GACCACTCCACTTATGGCTCTTTCGACTGCGCCAGTTTATGTTCTGACTCTGCTGACTTG 1500
DB      1441 GACCACTCCACTTATGGCTCTTTCGACTGCGCCAGTTTATGTTCTGACTCTGCTGACTTG 1500

QY      1501 GTTATGTTGCGACCGCGCGAGCGCTTGGCCGCGTCTCGATGGAACCAAGTCA 1560
DB      1501 GTTATGTTGCGACCGCGCGAGCGCTTGGCCGCGTCTCGATGGAACCAAGTCA 1560

QY      1561 CTTGAGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAGACCTTCTTGTCTGCGG 1620
DB      1561 CTTGAGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAGACCTTCTTGTCTGCGG 1620

QY      1621 CTCGCGGTAAGCTCTCTTTCGAGGAGCGAGCACAACTAAAGCCGGGTACCCCTTATAAT 1680
DB      1621 CTCGCGGTAAGCTCTCTTTCGAGGAGCGAGCACAACTAAAGCCGGGTACCCCTTATAAT 1680

QY      1681 TATAACCACTGCTAGCGACCAACTGTTGTCGAGGATGCGCGCGGACCGGGTCTGCT 1740
DB      1681 TATAACCACTGCTAGCGACCAACTGTTGTCGAGGATGCGCGCGGACCGGGTCTGCT 1740

QY      1741 ATTTCCACTTACACCACTAGCTGCTGCTGCTCCGCTCTCATTTCTGCGGTTGCCGTT 1800
DB      1741 ATTTCCACTTACACCACTAGCTGCTGCTGCTCCGCTCTCATTTCTGCGGTTGCCGTT 1800

QY      1801 TTAGCCCGCACTCTGGCTAGCAATGCTTGGAGTACCTTGGACTACCTTCCCGCGCC 1860
DB      1801 TTAGCCCGCACTCTGGCTAGCAATGCTTGGAGTACCTTGGACTACCTTCCCGCGCC 1860

QY      1861 CATACTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCCTTCAGGCTGCGCTTTC 1920
DB      1861 CATACTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCCTTCAGGCTGCGCTTTC 1920

QY      1921 CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTGAACTCGGAGTTG 1980
```

Db 1921 CAGTCTACTGTCGCTGAGCTTACGCGCCTTAGATGAGGTGGTAAACTCGGGAGTTG 1980
QY 1981 TAG 1983
Db 1981 TAG 1983
RESULT 2
US-10-165-868-3
; Sequence 3, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; Bradley, Daniel W.
; Twu, Jr-Shin
; Purdy, Michael A.
; Tam, Albert W.
; Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/165,868
; FILING DATE: 06-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 7
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
Us-10-165-868-3
Query Match 100.0%; Score 1983; DB 15; Length 2049;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCCTATTGTTGCTGCTCATGTTTGTGCTATGTCGCCGCGCCA 60
Db 1 ATGCGCCCTCGGCCTATTGTTGCTGCTCATGTTTGTGCTATGTCGCCGCGCCA 60
QY 61 CGCCCGGTCAGCGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 61 CGCCCGGTCAGCGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 121 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTTCGCAATCCCTATATTCATCCAAAC 180
Db 121 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTTCGCAATCCCTATATTCATCCAAAC 180
QY 181 CCCTTCGCCCGCGATGTCAACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db 181 CCCTTCGCCCGCGATGTCAACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 241 CGACCACTCGGCTCGGCTTGGGCGTGACAGGCGCGGCGGCGGCGGCGGCGGCGG 300
Db 241 CGACCACTCGGCTCGGCTTGGGCGTGACAGGCGCGGCGGCGGCGGCGGCGGCGG 300
QY 301 AGACCTACCAAGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Db 301 AGACCTACCAAGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 361 CAGTGCCTGATGTCGACTCCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Db 361 CAGTGCCTGATGTCGACTCCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 421 TCTCCCGCTTACCTCTTCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 480
Db 421 TCTCCCGCTTACCTCTTCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 480
QY 481 AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATATATGCGGCGGAGCTTCT 540
Db 481 AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATATATGCGGCGGAGCTTCT 540
QY 541 AATTATGCCAGTACCGGGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Db 541 AATTATGCCAGTACCGGGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
QY 601 GCTGTGCGGGTTACGGCATCTCCATCTCATCTGCGGCGGCGGCGGCGGCGGCGG 660
Db 601 GCTGTGCGGGTTACGGCATCTCCATCTCATCTGCGGCGGCGGCGGCGGCGGCGG 660
QY 661 TCGGTTGATATGAATTAATTAACCTCGACGGATGTTTCGTATTTAGTCCAGCGCGG 720
Db 661 TCGGTTGATATGAATTAATTAACCTCGACGGATGTTTCGTATTTAGTCCAGCGCGG 720
QY 721 GCTCTGAGCTTGTGATFCCAAAGTACGCGCTACATCTGTAACCAAGCGTGGCGTCC 780
Db 721 GCTCTGAGCTTGTGATFCCAAAGTACGCGCTACATCTGTAACCAAGCGTGGCGTCC 780
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCTGTTTGTATGCTTTGATA 840
Db 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCTGTTTGTATGCTTTGATA 840
QY 841 CATGGCTCACTCGTAAATTCCTATATACTATACACCTTATACCGGTGCCCTCGGGCTG 900
Db 841 CATGGCTCACTCGTAAATTCCTATATACTATACACCTTATACCGGTGCCCTCGGGCTG 900
QY 901 GACTTTGCCCTTGAGCTTGAGTTTTCGCAACCTTACCGCGGTAACCAATACGCGGTC 960
Db 901 GACTTTGCCCTTGAGCTTGAGTTTTCGCAACCTTACCGCGGTAACCAATACGCGGTC 960
QY 961 TCCGCTTATTCAGACACTGCTCGCCACCGCTTTCGTCGGGTCGGGAGCTGCGGAG 1020
Db 961 TCCGCTTATTCAGACACTGCTCGCCACCGCTTTCGTCGGGTCGGGAGCTGCGGAG 1020
QY 1021 CTCAACACCGGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGTT 1080
Db 1021 CTCAACACCGGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGTT 1080

Db 1021 CTCACCAACGCGTGTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080
QY 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACTCCCTGTTCAACCTTGTGACACTCTGCTT 1140
Db 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACTCCCTGTTCAACCTTGTGACACTCTGCTT 1140
QY 1141 GCGGCTCTGCGGACAGAAATGATTTGTCGCGTGTGTCGCGGAGCTGTTCTACTCCCGTCCC 1200
Db 1141 GCGGCTCTGCGGACAGAAATGATTTGTCGCGTGTGTCGCGGAGCTGTTCTACTCCCGTCCC 1200
QY 1201 GTTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATATCATCTGTAGAGAACTCTCAG 1260
Db 1201 GTTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATATCATCTGTAGAGAACTCTCAG 1260
QY 1261 CAGGATAGGCTATGCAATCCGCAATGACATGACCTCGGAGATCTCGTGTGTTATT 1320
Db 1261 CAGGATAGGCTATGCAATCCGCAATGACATGACCTCGGAGATCTCGTGTGTTATT 1320
QY 1321 CAGGATATGATAACCAACATGAACATGCGGCGAGCGCTTCTCCAGCCCGCATCGCGC 1380
Db 1321 CAGGATATGATAACCAACATGAACATGCGGCGAGCGCTTCTCCAGCCCGCATCGCGC 1380
QY 1381 CCTTCTCTGCTCTGAGCTAAATGATGTCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
Db 1381 CCTTCTCTGCTCTGAGCTAAATGATGTCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
QY 1441 GACGAGTCCACTTATGGCTCTTCACTGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 1441 GACGAGTCCACTTATGGCTCTTCACTGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
QY 1501 GTTAATGTTGCGACGCGCGCGAGCGCTTTCGCGGTGCTCGATGGAACCAAGTCA 1560
Db 1501 GTTAATGTTGCGACGCGCGCGAGCGCTTTCGCGGTGCTCGATGGAACCAAGTCA 1560
QY 1561 CTTGACGCTGCGCCCTCTCACCATCCAGAGTACTCGAAGACCTTCTTTGTCTGCGG 1620
Db 1561 CTTGACGCTGCGCCCTCTCACCATCCAGAGTACTCGAAGACCTTCTTTGTCTGCGG 1620
QY 1621 CTCGCGGTAAAGCTCTCTTCTGGGAGCGAGCAACCTAAAGCGGGTACCTTATAAT 1680
Db 1621 CTCGCGGTAAAGCTCTCTTCTGGGAGCGAGCAACCTAAAGCGGGTACCTTATAAT 1680
QY 1681 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAAATCCGCGCGGACCGGGTGCCT 1740
Db 1681 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAAATCCGCGCGGACCGGGTGCCT 1740
QY 1741 ATTTCCACTTACACACTAGCTGCTGGTGTGCTGCTCCGCTCCATTTCTGCGGTTGCCGTT 1800
Db 1741 ATTTCCACTTACACACTAGCTGCTGGTGTGCTGCTCCGCTCCATTTCTGCGGTTGCCGTT 1800
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTGCGCGGCC 1860
Db 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTGCGCGGCC 1860
QY 1861 CATACTTTTGTATGATTTCTGCGGAGTCCGCCCCCTTGGCCCTTTCAGGGTGTGCGTTTC 1920
Db 1861 CATACTTTTGTATGATTTCTGCGGAGTCCGCCCCCTTGGCCCTTTCAGGGTGTGCGTTTC 1920
QY 1921 CAGTCTACTGCTGAGCTTACGGCTTCAAGTGAAGTGGTAAAGTGGAGTTG 1980
Db 1921 CAGTCTACTGCTGAGCTTACGGCTTCAAGTGAAGTGGTAAAGTGGAGTTG 1980
QY 1981 TAG 1983
Db 1981 TAG 1983

RESULT 3

US-10-165-868-1

; Sequence 1, Application US/10165868

; Publication No. US20030143241A1

; GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory R.

Bradley, Daniel W.
Twu, Jr-Shin
Purdy, Michael A.
Tam, Albert W.
Krawczynski, Krzysztof Z.
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/165,868
FILING DATE: 06-Jun-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/870,985A
FILING DATE: 20-APRIL-1992
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0093.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2094 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: BURMA SEQUENCE, FIGURE 7
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-165-868-1

Query Match 100.0%; Score 1983; DB 15; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCGGCTATTTTGTCTCTCTCATGTTTGTCTCTATGCTCGCGCGGCA 60
Db 46 ATGCGCCCTCGGCTATTTTGTCTCTCTCATGTTTGTCTCTATGCTCGCGCGGCA 105
QY 61 CCGCCCGGTGAGCCGTCTGCGCGCCGCTGCTGGGCGGCGAGCGCGGTTCCGCGGTGGT 120
Db 106 CCGCCCGGTGAGCCGTCTGCGCGCCGCTGCTGGGCGGCGAGCGCGGTTCCGCGGTGGT 165
QY 121 TTCTGGGGTGACCGGGTTGATTTCTCAGCCCTTCCCAATCCCTATATTTCATCCACCAAC 180
Db 166 TTCTGGGGTGACCGGGTTGATTTCTCAGCCCTTCCCAATCCCTATATTTCATCCACCAAC 225

181 CCCTTCGCCCCGATGTCACCGCTGCGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCG 240
182 |||||
226 CCCTTCGCCCCGATGTCACCGCTGCGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCG 285
227 |||||
241 CGACCACTCGGCTCCGCTTGGGCTGACCGAGCCACGCGCCGCGCTTGCCTCACGTCGT 300
242 |||||
286 CGACCACTCGGCTCCGCTTGGGCTGACCGAGCCACGCGCCGCGCTTGCCTCACGTCGT 345
287 |||||
301 AGACCTACACAGCTGGGCGCGCGCTTAACCGCGGTGCTCCGCGCCATGACACCCGCG 360
302 |||||
346 AGACCTACACAGCTGGGCGCGCGCTTAACCGCGGTGCTCCGCGCCATGACACCCGCG 405
347 |||||
361 CCAGTGCCTGATGTCGACTCCGCGCGCGCATCTTCGCGCGGCGAGTAACTATCAACA 420
362 |||||
406 CCAGTGCCTGATGTCGACTCCGCGCGGCGCATCTTCGCGCGGCGAGTAACTATCAACA 465
407 |||||
421 TCTCCCTTACCTTTCGCTGGCCACCGGCACTAACCTGGTCTCTTTATGCGCGCCCTCTT 480
422 |||||
466 TCTCCCTTACCTTTCGCTGGCCACCGGCACTAACCTGGTCTCTTTATGCGCGCCCTCTT 525
467 |||||
481 AGTCCGCTTTTACCCCTTCAGACGCGACCAATACCCATATATGCGCGAGGAGTTCT 540
482 |||||
526 AGTCCGCTTTTACCCCTTCAGACGCGACCAATACCCATATATGCGCGAGGAGTTCT 585
527 |||||
541 AATTATGCCAGTACCGGCTTGGCGGTGCGCAACATCCGTTACCGCGCGCTGGTCCCAAT 600
542 |||||
586 AATTATGCCAGTACCGGCTTGGCGGTGCGCAACATCCGTTACCGCGCGCTGGTCCCAAT 645
587 |||||
601 GCTGTCGGGGTTAGCCCATCTCCATCTCATCTTGGGCCACAGACCAACCCACCGGAGC 660
602 |||||
646 GCTGTCGGGGTTAGCCCATCTCCATCTCATCTTGGGCCACAGACCAACCCACCGGAGC 705
647 |||||
661 TCCGTTGATATCAATTAATACCTCGACGAGTTCGTAATTTAGTCAGGCCCGGCATA 720
662 |||||
706 TCCGTTGATATCAATTAATACCTCGACGAGTTCGTAATTTAGTCAGGCCCGGCATA 765
707 |||||
721 GCCTCTGAGCTTGTGATCTCCAAAGTAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
722 |||||
766 GCCTCTGAGCTTGTGATCCCAAGTAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 825
767 |||||
781 GTCGAGACTCTGGGGTGGCTGAGGAGGAGTACCTCTGCTGTTGTTATGCTTTGCATA 840
782 |||||
826 GTCGAGACTCTGGGGTGGCTGAGGAGGAGTACCTCTGCTGTTGTTATGCTTTGCATA 885
827 |||||
841 CATGCTCACTCGTAAATTCCTATACTAATACCCCTATACCGGTGCGCTCGGGCTGTTG 900
842 |||||
886 CATGCTCACTCGTAAATTCCTATACTAATACCCCTATACCGGTGCGCTCGGGCTGTTG 945
887 |||||
901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATAACGCGGTC 960
902 |||||
946 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATAACGCGGTC 1005
947 |||||
961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTCGCGTGGCGGAGCGGACTGCCGAG 1020
962 |||||
1006 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTCGCGTGGCGGAGCGGACTGCCGAG 1065
1007 |||||
1021 CTCACCAACGCGCTGCTACCGCTTTATGAGGACCTCTATTTACTAGTACTAATGTT 1080
1022 |||||
1066 CTCACCAACGCGCTGCTACCGCTTTATGAGGACCTCTATTTACTAGTACTAATGTT 1125
1067 |||||
1081 GTCGCTGAGATCGGCGCGGAGTACCTCTCACCTGTTTCAACCTTGTGACACTCTGCTT 1140
1082 |||||
1126 GTCGCTGAGATCGGCGCGGAGTACCTCTCACCTGTTTCAACCTTGTGACACTCTGCTT 1185
1127 |||||
1141 GCGCGCTTCGCGACAGAAATGATTTGTCGCGTGGTGCCAGCTGTTCTACTCCGCTCC 1200
1142 |||||
1186 GCGCGCTTCGCGACAGAAATGATTTGTCGCGTGGTGCCAGCTGTTCTACTCCGCTCC 1245
1187 |||||
1201 GTTGTCTCAGCAATGGCGAGCGGAGCTGTTAAGTTGTATACATCTGTAGAGAACTCTCAG 1260
1202 |||||
1246 GTTGTCTCAGCAATGGCGAGCGGAGCTGTTAAGTTGTATACATCTGTAGAGAACTCTCAG 1305
1247 |||||

1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAACTCGTGTGGTTATT 1320
1262 |||||
1306 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAACTCGTGTGGTTATT 1365
1307 |||||
1321 CAGGATTATGATAACCAACCAATGAAACAAAGATCGGCGCGGCTTCTCCAGCCCCATCGCGC 1380
1322 |||||
1366 CAGGATTATGATAACCAACCAATGAAACAAAGATCGGCGCGGCTTCTCCAGCCCCATCGCGC 1425
1367 |||||
1381 CTTTCTCTGTCTCTGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
1382 |||||
1426 CTTTCTCTGTCTCTGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1485
1427 |||||
1441 GACCACTGACATGAGCTCTCTCTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
1442 |||||
1486 GACCACTGACATGAGCTCTCTCTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1545
1487 |||||
1501 GTTAATGTTGCGACCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
1502 |||||
1546 GTTAATGTTGCGACCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1605
1547 |||||
1561 CTTGAGCGTTCGCGCGCTCTCTCAACATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGCG 1620
1562 |||||
1606 CTTGAGCGTTCGCGCGCTCTCTCAACATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGCG 1665
1607 |||||
1621 CTCGCGGTAAAGCTCTCTCTCTGAGGAGCAGGCAAACTAAAGCGGGTACCTTTATAAT 1680
1622 |||||
1666 CTCGCGGTAAAGCTCTCTCTCTGAGGAGCAGGCAAACTAAAGCGGGTACCTTTATAAT 1725
1667 |||||
1681 TATAACACCACTGTAGCGACCAACTGCTTCTGAGAAATGCGCGCGGCGCACCGGGTGGCT 1740
1682 |||||
1726 TATAACACCACTGTAGCGACCAACTGCTTCTGAGAAATGCGCGCGGCGCACCGGGTGGCT 1785
1727 |||||
1741 ATTTCACCTTACACCACTGTAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1742 |||||
1786 ATTTCACCTTACACCACTGTAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1845
1787 |||||
1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTTGGCGCGCGC 1860
1802 |||||
1846 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTTGGCGCGCGC 1905
1847 |||||
1861 CATACTTTTGTATGATTTCTGCGCGAGTTCGCGCGCGCTTGGCGCTTTCAGGGCTGCGCTTTC 1920
1862 |||||
1906 CATACTTTTGTATGATTTCTGCGCGAGTTCGCGCGCGCTTGGCGCTTTCAGGGCTGCGCTTTC 1965
1907 |||||
1921 CAGTCTACTGCTGCTGAGCTTACGCGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 1980
1922 |||||
1966 CAGTCTACTGCTGCTGAGCTTACGCGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 2025
1967 |||||
1981 TAG 1983
1982 |||||
2026 TAG 2028

RESULT 4

US-09-851-410-6

; Sequence 6, Application US/09851410

; Publication No. US20030124510A1

; GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory R

; Yarbrough, Patrice O

; Bradley, Daniel W

; Krawczynski, Krzysztof Z

; Tam, Albert

; Fry, Kirk E

; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted

; No. US20030124510A1-A/No. US20030124510A1-B Hepatitis Viral

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates

; STREET: 350 Cambridge Avenue, Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

1141 GCGCGCTCCGACAGAAATGATTTCTGCGCTGGTGGCCAGCTGTTCTACTCCGTCCTC 1200
1141 GCGCGCTCCGACAGAAATGATTTCTGCGCTGGTGGCCAGCTGTTCTACTCCGTCCTC 1200
1201 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGATATACATCTGTAGAGAACTCAG 1260
1201 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGATATACATCTGTAGAGAACTCAG 1260
1261 CAGGATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1261 CAGGATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1321 CAGGATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1321 CAGGATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1381 CTTTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1440
1381 CTTTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1440
1441 GACCACTCCACTTATGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1500
1441 GACCACTCCACTTATGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1500
1501 GTTAATGTTGCGACCGCGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
1501 GTTAATGTTGCGACCGCGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
1561 CTTGAGCGTGCCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1620
1561 CTTGAGCGTGCCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1620
1621 CTTGAGCGTGCCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1680
1621 CTTGAGCGTGCCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1680
1681 TATAACCACTGCTAGTGAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
1681 TATAACCACTGCTAGTGAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
1741 ATTTCCACTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1741 ATTTCCACTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTTGAGGATACCTTGAAGTACCTGCGGCGCC 1860
1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTTGAGGATACCTTGAAGTACCTGCGGCGCC 1860
1861 CATACTTTGATGATTTCTGCGAGAGTGGCGGCGGCTTGGCTTCAAGGCTGCGCTTTC 1920
1861 CATACTTTGATGATTTCTGCGAGAGTGGCGGCGGCTTGGCTTCAAGGCTGCGCTTTC 1920
1921 CAGTCTACTGCTGCTGAGCTTCAAGGCTTGAAGTGAAGTGGTGAAGTGAAGTGAAGTGA 1980
1921 CAGTCTACTGCTGCTGAGCTTCAAGGCTTGAAGTGAAGTGGTGAAGTGAAGTGAAGTGA 1980
1981 TAG 1983
1981 TAG 1983

Query Match 90.4%; Score 1793.4; DB 13; Length 1989;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 1871; Conservative 0; Mismatches 111; Indels 1; Gaps 1;
US-10-381-770-6
1 ATGCGCCCTCGGCTTATTTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 60
1 ATGCGCCCTCGGCTTATTTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 60
61 CCGCCCGGTGACCGCTCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCTGCTGCGCGCGCT 120
61 CCGCCCGGTGACCGCTCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCTGCTGCGCGCGCT 120
121 TTCTGCGGTGACCGCGCTTGAATCTCAGCCCTTTCGCAATCCCTATATTCATCAACCAAC 180
121 TTCTGCGGTGACCGCGCTTGAATCTCAGCCCTTTCGCAATCCCTATATTCATCAACCAAC 180
181 CCGCTTCGCGCGCGCTGCTGCGCGCGCGCTGCTGCGCGCGCGCTGCTGCGCGCGCGCT 240
181 CCGCTTCGCGCGCGCTGCTGCGCGCGCGCTGCTGCGCGCGCGCTGCTGCGCGCGCGCT 240
241 CGACCACTCGGCTTCCGCTTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 300
241 CGACCACTCGGCTTCCGCTTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 300
301 AGACCTTACCACTGCGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCTGCTGCGCGCGCT 360
301 AGACCTTACCACTGCGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCTGCTGCGCGCGCT 360
361 CAGTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
361 CAGTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
421 TCTCCCTTACCTTCCGCTGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCTGCTGCGCGCT 480
421 TCTCCCTTACCTTCCGCTGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCTGCTGCGCGCT 480
481 AGTCCGCTTTCACCTTTCAGGACCGGCAACCAATACCAATATATATGCGCGCGCTTCT 540
481 AGTCCGCTTTCACCTTTCAGGACCGGCAACCAATACCAATATATATGCGCGCGCTTCT 540
541 AATATGCGCGCTTACCGGCTTCCGCTGCGCGCTTAAACCGCGCTGCTGCGCGCGCTTCT 600
541 AATATGCGCGCTTACCGGCTTCCGCTGCGCGCTTAAACCGCGCTGCTGCGCGCGCTTCT 600
601 GCTGCGCGCTTACCGGCTTCCGCTGCGCGCTTAAACCGCGCTGCTGCGCGCGCTTCT 660
601 GCTGCGCGCTTACCGGCTTCCGCTGCGCGCTTAAACCGCGCTGCTGCGCGCGCTTCT 660
661 TCCGCTTGCATGATGATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 720
661 TCCGCTTGCATGATGATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 720
721 GCCTCTGAGCTTGTGATCCCAAGTGGCGCTTCACTATCGTAACCAAGGCTGCGGCTTCT 780
721 GCCTCTGAGCTTGTGATCCCAAGTGGCGCTTCACTATCGTAACCAAGGCTGCGGCTTCT 780
781 GTCGAGACTCTCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
781 GTCGAGACTCTCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
841 CATGCTCACTGCTGATAATTCATTAATCACTATACCGGCTGCGGCTGCGGCTGCTGCTGCT 900
841 CATGCTCACTGCTGATAATTCATTAATCACTATACCGGCTGCGGCTGCGGCTGCTGCTGCT 900
901 GACITTTGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 960

RESULT 6
US-10-381-770-6
Sequence 6, Application US/10381770
Publication No. US20040052813A1
GENERAL INFORMATION:
APPLICANT: YANG SHENG TANG company, Ltd.
TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS THE VACCINE COMPOS
TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
FILE REFERENCE: IEC010037PCT
CURRENT APPLICATION NUMBER: US/10/381,770
CURRENT FILING DATE: 2003-07-16

Db 901 GACTTTGCGCTCGAATGTAGTTCCGCAACCTCACCCCGGTATATACCAACGCGGGTC 960
QY 961 TCCGTTTATTCAGACACTGCTCGCACCGCTTCGTGCGGTCGGAGCGGACTGCGGAG 1020
Db 961 TCCGTTTATTCAGACACTGCTCGCACCGCTTCGTGCGGTCGGAGCGGACTGCGGAG 1020
QY 1021 CTCACACACGCGCTGTACCCGCTTTATGAAGGACCTCTATTTACTAGTACTAATGGT 1080
Db 1021 CTTTACCACACGCGCTGTACCCGCTTTATGAAGGACCTCTATTTACTAGTACTAATGGT 1080
QY 1081 GTCCGTTAGATCGCGCGCGGATAGCCCTCACCTGTTCAACCTTGTGCTGACACTCTGCTT 1140
Db 1081 GTCCGTTAGATCGCGCGCGGATAGCCCTCACCTGTTCAACCTTGTGCTGACACTCTGCTT 1140
QY 1141 GCGCGCTCCGACAGAAATGATTTGCTGCGCTGGTGGCGAGCTGTTCTACTCCGCTGCC 1200
Db 1141 GCGCGCTTACCACAGAAATGATTTGCTGCGCTGGTGGCGAGCTGTTCTACTCTCGTCCC 1200
QY 1201 GTTGTCTAGCAATGCGCGAGCGGACTGTTAGTTGTATATCTGTAGAGAAATGCTCAG 1260
Db 1201 GTTGTCTAGCAATGCGCGAGCGGACTGTTAGTTGTATATCTGTAGAGAAATGCTCAG 1260
QY 1261 CAGGATTAAGGTTATGCAATCCGCGATGACATGACCTCGGAAATCTGTTGTTGTTATT 1320
Db 1261 CAGGATTAAGGTTATGCAATCCGCGATGACATGACCTCGGAGGAGTCTGTTGTTATT 1320
QY 1321 CAGGATTAAGTAAACCAACATGAACAGATCGCGCGGAGCTTCTCCAGCCCATCGCGC 1380
Db 1321 CAGGATTAAGTAAACCAACATGACGAGCAGCAGCAGCTTCCAGCCCATCGCGC 1380
QY 1381 CTTTCTCTGCTCTGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTGCGAGTAT 1440
Db 1381 CTTTCTCTGCTCTGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTGCGAGTAT 1440
QY 1441 GACAGTCCACTTATGCTCTGAGCTGCGGCGGCTTATGTTCTGACTCTGTGACCTTG 1500
Db 1441 GACAGTCCACTTATGCTCTGAGCTGCGGCGGCTTATGTTCTGACTCTGTGACCTTG 1500
QY 1501 GTTAATGTTGGACCGCGCGGAGCGGCTTGGCGGCTCGCTCGATGGACCAAGGTACA 1560
Db 1501 GTTAATGTTGGACCGCGCGGAGCGGCTTGGCGGCTCGCTCGATGGACCAAGGTACA 1560
QY 1561 CTTGACGCTGCGGCGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGG 1620
Db 1561 CTTGATGCTGCGGCGCTTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGG 1620
QY 1621 CTCGCGGTAAAGCTCTCTTTCTGGGAGCAGCACAATAAGCGGTACCTTATTAAT 1680
Db 1621 CTCGCGGTAAAGCTCTCTTTGGAGCAGGTACTACTAAGCGGCTTACCTTATAAT 1680
QY 1681 TATAACACCACTGTAGCAGCAACTGCTTGTGAGAAATGCGCGGCGACCGGCTGCT 1740
Db 1681 TATAACACCACTGTAGCAGCAACTGCTTGTGAGAAATGCGCGGCGACCGGCTGCT 1740
QY 1741 ATTTCCCACTTACACACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1741 ATTTCCCACTTACACACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
QY 1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTGAAGATACCTTGAATACCTTGCCTGCGG 1860
Db 1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTGAAGATACCTTGAATACCTTGCCTGCGG 1860
QY 1861 CATATTTTGTATGATTTTGTCCAGAGTGGCGGCTTGGCTTTCAGGGCTGCGCTTTC 1920
Db 1860 CATATTTTGTATGATTTTGTCCAGAGTGGCGGCTTGGCTTTCAGGGCTGCGCTTTC 1920
QY 1921 CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGTAAAGTGGGTAATACTCGGAGTTG 1980
Db 1920 CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGTAAAGTGGGTAATACTCGGAGTTA 1980
QY 1981 TAG 1983

Db 1980 TAG 1982
RESULT 7
US-09-769-066-3
; Sequence 3, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; McAttee, C. Patrick
; Yarbrough, Patrice O.
; Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/769,066
; FILING DATE: 24-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/542,634
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1647 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E virus (Burma) r62kDa,
; FIGURE 2
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-769-066-3

Query Match 83.1%; Score 1647; DB 9; Length 1647;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 GCGGTCTCGGCGCCATGACACCCGCGCAGTGCCTGATGCTGACTCCCGGGCGGCATC 393
Db 1 GCGGTCTCGGCGCCATGACACCCGCGCAGTGCCTGATGCTGACTCCCGGGCGGCATC 60
QY 394 TTGGCGCGGAGTATAACCTTATCAACATCTCCCTTTACCTTCTCCGTTGGCCACCGGCACT 453
Db 61 TTGGCGCGGAGTATAACCTTATCAACATCTCCCTTTACCTTCTCCGTTGGCCACCGGCACT 120
QY 454 AACCTGGTCTTTTATGCGGCGCCCTCTTAGTCCGCTTTTACCCCTTCAGGACGCAACAAT 513
Db 121 AACCTGGTCTTTTATGCGGCGCCCTCTTAGTCCGCTTTTACCCCTTCAGGACGCAACAAT 180
QY 514 ACCCATATATGCGCACGGAAGCTTCTAATTATGCCCCAGTACCGGGTTGCCCGTGCCACA 573
Db 181 ACCCATATATGCGCACGGAAGCTTCTAATTATGCCCCAGTACCGGGTTGCCCGTGCCACA 240

QY	1654	ACAACTAAAGCCGGGTACCTCTAATATTAAACACCACTCTAGTCGACCAACTGCTTGTC	1711
Db	1321	ACAACTAAAGCCGGGTACCTCTAATATTAAACACCACTCTAGTCGACCAACTGCTTGTC	1380
QY	1714	GAGAACTGCCCGCCGGGCACCGGGTCGCTAATTTCCACTTACACCACTAGCTGGGTGCTGGT	1773
Db	1381	GAGAACTGCCCGCCGGGCACCGGGTCGCTAATTTCCACTTACACCACTAGCTGGGTGCTGGT	1440
QY	1774	CCCGTCTCCAAATTCCTGCGGTTGCGGTTTATAGCCCCCACCCTCTGCGCTAGCAATTGCTTTGAG	1833
Db	1441	CCCGTCTCCAAATTCCTGCGGTTGCGGTTTATAGCCCCCACCCTCTGCGCTAGCAATTGCTTTGAG	1500
QY	1834	GATACTTTGGACTACCTGCGCGGCCCACTATTTTGATTTCTGCCCGCAGAGTGCCGC	1893
Db	1501	GATACTTTGGACTACCTGCGCGGCCCACTATTTTGATTTCTGCCCGCAGAGTGCCGC	1560
QY	1894	CCCGTTGGCCTTTGAGGCTCGGCTTTCCAGTCTACTGCTGAGCTTCAGCGGCTTTAAG	1953
Db	1561	CCCGTTGGCCTTTGAGGCTCGGCTTTCCAGTCTACTGCTGAGCTTCAGCGGCTTTAAG	1620
QY	1954	ATGAAGGTGGGTAAAACTCGGGAGTTG	1980
Db	1621	ATGAAGGTGGGTAAAACTCGGGAGTTG	1647

RESULT 8

US-09-769-066-2
; Sequence 2, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; McAtee, C. Patricia
; Yarbough, Patrice O.
; Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/769,066
; FILING DATE: 24-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/542,634
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2058 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: Hepatitis E Virus (Mexico Strain)
; ORF-2 region
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

QY	574	ATCCGTTACGCCCGCGCTCGTCCCCAAATGCTGTCGGCGGTACGCCAATCTCCATCTCATTC	633
DB	241	ATCCGTTACGCCCGCGCTCGTCCCCAAATGCTGTCGGCGGTACGCCAATCTCCATCTCATTC	300
QY	634	TGGCCACAGACACACACACCCCGAGCTCCGTTGATGAATGAATCAATAACCTCGAGCGAT	693
DB	301	TGGCCACAGACACACACACCCCGAGCTCCGTTGATGAATGAATCAATAACCTCGAGCGAT	360
QY	694	GTTTCGTATTATTAGTTCAGCGCCGCGATAGCCTCTGAGCTTTGTGATCCCAAGTAGCGGCTA	753
DB	361	GTTTCGTATTATTAGTTCAGCGCCGCGATAGCCTCTGAGCTTTGTGATCCCAAGTAGCGGCTA	420
QY	754	CACATATCGTAACCAAGGCTTGGCGCTCCGTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCT	813
DB	421	CACATATCGTAACCAAGGCTTGGCGCTCCGTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCT	480
QY	814	ACCTCTGGTCTTGTATTATGCTTTGCATACATGCGCTCACTCGTAAATTCCTATACPAATA	873
DB	481	ACCTCTGGTCTTGTATTATGCTTTGCATACATGCGCTCACTCGTAAATTCCTATACPAATA	540
QY	874	CCCTATACCGGTGCCCTTCGGGCTTTTGGACTTTTGGCCCTTGAGCTTTGAGTTTCGCAACCTT	933
DB	541	CCCTATACCGGTGCCCTTCGGGCTTTTGGACTTTTGGCCCTTGAGCTTTGAGTTTCGCAACCTT	600
QY	934	ACCCCGGTAAACCAATACCGCGTCTCCGTTTATTCACGACCTGCTCGCCACACCGCTT	993
DB	601	ACCCCGGTAAACCAATACCGCGTCTCCGTTTATTCACGACCTGCTCGCCACACCGCTT	660
QY	994	CGTTCGGGTTCGGAACGGGACTGCGAGCTCACCAACGCGTGTACCCGCTTTATGAG	1053
DB	661	CGTTCGGGTTCGGAACGGGACTGCGAGCTCACCAACGCGTGTACCCGCTTTATGAG	720
QY	1054	GACCTCTATTTTACTAGTACTAATGTTGCTCGGTGAGATCGGCGCGGATAGCCCTCACC	1113
DB	721	GACCTCTATTTTACTAGTACTAATGTTGCTCGGTGAGATCGGCGCGGATAGCCCTCACC	780
QY	1114	CTGTTCAACCTTGTCTGACACTCTGTTTGGCGGCGCTGCCGACAGAAATTTGTCGCGCT	1173
DB	781	CTGTTCAACCTTGTCTGACACTCTGTTTGGCGGCGCTGCCGACAGAAATTTGTCGCGCT	840
QY	1174	GGTGGCGAGCTGTTCTACTCCCGTCCCGTGTCTCAGCCAATGGCGAGCCGACTGTTAAG	1233
DB	841	GGTGGCGAGCTGTTCTACTCCCGTCCCGTGTCTCAGCCAATGGCGAGCCGACTGTTAAG	900
QY	1234	TTGTATACATCTGTAGAGAAATGCTCAGCAGGATAGGGTATTGCCAATCCCGATGACATT	1293
DB	901	TTGTATACATCTGTAGAGAAATGCTCAGCAGGATAGGGTATTGCCAATCCCGATGACATT	960
QY	1294	GACCTCGGAGAAATCTCGTGTGGTTATTGAGGTTATGATACCAATGAAACAGATCGG	1353
DB	961	GACCTCGGAGAAATCTCGTGTGGTTATTGAGGTTATGATACCAATGAAACAGATCGG	1020
QY	1354	CCGAGCGCTTCTCCAGCCCGCATCGCGCCCTTTCTCTGTCTTTCGAGCTAATGATGCTT	1413
DB	1021	CCGAGCGCTTCTCCAGCCCGCATCGCGCCCTTTCTCTGTCTTTCGAGCTAATGATGCTT	1080
QY	1414	TGGTCTCTCTCACCGTCCGAGTATGACAGTCCATTATGGCTCTTTCGAGTCGCCCA	1473
DB	1081	TGGTCTCTCTCACCGTCCGAGTATGACAGTCCATTATGGCTCTTTCGAGTCGCCCA	1140
QY	1474	GTTTATGTTTCTGACTCTGTGACCTTGGTTAATGTTGACCGCGCGCAGGCGCTTGCC	1533
DB	1141	GTTTATGTTTCTGACTCTGTGACCTTGGTTAATGTTGACCGCGCGCAGGCGCTTGCC	1200
QY	1534	CGGTTCGCTCGATTGGACCAAGGTCACATGACGCTCGGCCCTCTTCGACCATCCAGCAG	1593
DB	1201	CGGTTCGCTCGATTGGACCAAGGTCACATGACGCTCGGCCCTCTTCGACCATCCAGCAG	1260
QY	1594	TACTCGAGACCTTCTTTTGTCTCGCGCTCCGCGGTAAAGCTCTCTTTTCTGGGAGGAGGC	1653
DB	1261	TACTCGAGACCTTCTTTTGTCTCGCGCTCCGCGGTAAAGCTCTCTTTTCTGGGAGGAGGC	1320

Publication No. US20030143241A1	
GENERAL INFORMATION:	
APPLICANT: Reyes, Gregory R.	
Bradley, Daniel W.	
Twu, Jr-Shin	
Purdy, Michael A.	
Tam, Albert W.	
Krawczynski, Krzysztof Z.	
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method	
NUMBER OF SEQUENCES: 22	
CORRESPONDENCE ADDRESS:	
ADDRESSER: Dehlinger & Associates	
STREET: 350 Cambridge Avenue, Suite 250	
CITY: Palo Alto	
STATE: CA	
COUNTRY: USA	
ZIP: 94306	
COMPUTER READABLE FORM:	
MEDIUM TYPE: Floppy disk	
COMPUTER: IBM PC compatible	
OPERATING SYSTEM: PC-DOS/MS-DOS	
SOFTWARE: Patent in Release #1.0, Version #1.25	
CURRENT APPLICATION DATA:	
APPLICATION NUMBER: US/10/165,868	
FILING DATE: 06-Jun-2002	
CLASSIFICATION: <Unknown>	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: US/07/870,985A	
FILING DATE: 20-APRIL-1992	
APPLICATION NUMBER: US 822,335	
FILING DATE: 17-JAN-1992	
APPLICATION NUMBER: US 505,888	
FILING DATE: 05-APRIL-1990	
APPLICATION NUMBER: US 420,921	
FILING DATE: 13-OCTOBER-1989	
APPLICATION NUMBER: US 367,486	
FILING DATE: 16-JUNE-1989	
APPLICATION NUMBER: US 336,672	
FILING DATE: 11-APRIL-1989	
APPLICATION NUMBER: US 208,997	
FILING DATE: 17-JUNE-1988	
ATTORNEY/AGENT INFORMATION:	
NAME: Fabian, Gary R.	
REGISTRATION NUMBER: 33,875	
REFERENCE/DOCKET NUMBER: 4600-0093.30	
TELECOMMUNICATION INFORMATION:	
TELEPHONE: (415) 324-0880	
INFORMATION FOR SEQ ID NO: 4:	
SEQUENCE CHARACTERISTICS:	
LENGTH: 2055 base pairs	
TYPE: nucleic acid	
STRANDEDNESS: double	
TOPOLOGY: linear	
MOLECULE TYPE: cdna to mRNA	
HYPOTHETICAL: NO	
ORIGINAL SOURCE:	
INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 7	
SEQUENCE DESCRIPTION: SEQ ID NO: 4:	
US-10-165-868-4	
Query Match	69.0%; Score 1368.6; DB 15; Length 2055;
Best Local Similarity	81.0%; Pred. No. 0;
Matches 1606; Conservative	0; Mismatches 374; Indels 3; Gaps 1;
QY	1 ATGCGCCCTCGCCCTATTTTGTGCTGCTCTCATGTTTTTGGCCCTATGCTCGCCGCGCA 60
DB	1 ATGCGCCCTTAGCGCCCTTTTGTGTTTCTCTTGTGTTTCTGCTATGTTTGGCCGCGCA 60
QY	61 CCGCCCGGTTCAGCCGCTCTGCGCCGCGCTGTCGGGCGCGCAGCGCGGTTCGCGCGGTGGT 120
DB	61 CCGACCGGTTCAGCCGCTCTGCGCCGCGCTGTCGGGCGCGCAGCGCGGTTCAGCCGCGGTGGT 120

Db 346 CGACCTGCCACGCGCGGCTCGCGGCTGACGGCTGTGGCGCTGCCCATGACACTCA 405
Qy 361 CCAGTGCCTGATGTCGACTCCCGCGCGCAATCTTTGGCGCGGCACTAATCAACA 420
Db 406 CCCGTCGGGAGGTTGATTTCTCGCGGTGCAATTTCTACGCGCGCAGTAAATTTGTCTACT 465
Qy 421 TCTCCCTTACCTCTTCGTCGCGGACCGGCACTAATCTGTTCTTTATGCGCGCCCTCTT 480
Db 466 TCACCCCTGACATCTCTGTGGCTCTGGCACTAATTTAGTCTCTGTATGACGCCCTT 525
Qy 481 AGTCGCTTTTACCCCTTCAGGACGCGCACCAATACCATATATATGGCCACGGAAGCTTCT 540
Db 526 AATTCGCTCTGCGGTGAGGACGGTACTAATACTCAATATATGGCCACAGAGGCTCC 585
Qy 541 AATATATGCCAGTACCGGTTGCGGTGCGCAAACTCCGTTTACCGCCCGTGTGCCCAAT 600
Db 586 AATATGACAGTACCGGTTGCGCGCTACTATCCGTTTACCGGCCCTTAGTGCCCTAAT 645
Qy 601 GCTGTGCGGCTTACGCCATCTCCATCTCAATCTGCGCACAGACCAACCAACCCCTTCA 660
Db 646 GCAGTTGGAGGCTATGCTATATCCATTTCTTTCTGGCCTCAAAACAACCAACCCCTTCA 705
Qy 661 TCGGTTGATGATGAATTCATTAACCTCGAGGATGTTCTGATTTTGTAGTCCAGCCCGCAT 720
Db 706 TCTGTTGACATGATTTCCATTAATCTTCCATGATGTCAGGATTTCTGTTCAACCTGCA 765
Qy 721 GCTCTGAGCTTGTGATCCCAAGTACGCGCTACATATGCTGTAACAGGCTGGCGCTCC 780
Db 766 GCATCTGAATGCTCATCCCAAGGAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 825
Qy 781 GTGAGACCTCTGGGTGCTGAGGAGGCTACCTCTGCTCTGTTGATGTTTGCATA 840
Db 826 GTTGAGACATCTGGTGTGCTGAGGAGGAGCACTCCGCTCTGTCATGTTATGATA 885
Qy 841 CATGGCTCACTCCGTAATTCCTATCTAATAACCTATACCCCTTATACCGGTGCGGTGTTG 900
Db 886 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCGGTGTTGTT 945
Qy 901 GACTTTGCCCTTGAGCTTGATTTGCAACCTTACCCCGGTAAACCAATACCGGCTC 960
Db 946 GACTTTGCCCTTAGAGCTTGATTTGCAATCTCACCACCTGTAAACCAATACCGGTG 1005
Qy 961 TCCGCTTATTCAGCACTGCTCGCCACCGCTTCTGTCGCGGTGCGGACGGAAGTCCGAG 1020
Db 1006 TCCGCTTATTCAGCACTGCTGTCATCC---GCCGAGGCGGACGAGCTGCGGAG 1062
Qy 1021 CTCACCAACCGGTGCTACCGCTTTATGAAGGACCTCTATTTACTAGTACTAATGTT 1080
Db 1063 CTGACCACTGACGAGCAGCAGGTTCTATGAAGATCTCCACTTTACCGGCTTAATGGG 1122
Qy 1081 GTCGCTGAGATCGGCGGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTCTGCTT 1140
Db 1123 GTAGGTGAAGTTCGCGCGGCGGATAGCTCTAACATTTACTTAACCTTGTGACAGCTCCTC 1182
Qy 1141 GCGGCTCGCGCAGCAATGATTTGTCGCTGCTGCGCAGCTGTTCTACTCCCGTCCC 1200
Db 1183 GCGGCTCGCGCAGCAATTAATTTGTCGCTGCGCGGCACTGTTTATTCGCGCGG 1242
Qy 1201 GTTCTCTCAGCAATGCGGCGGAGCTGTTAAGTTGTATACATCTGAGAGAATGCTCAG 1260
Db 1243 GTTCTCTCAGCAATGCGGCGGAGCAGCTGAAGCTCTATACATCAGTGGAGAATGCTCAG 1302
Qy 1261 CAGATTAAGGTTATGCAATCCGCGAGTACATGACCTCGAGAATCTCGTGTGTTATT 1320
Db 1303 CAGGATAAGGTTGTTGATATCCCGCAGATATCATCTTGTGATTTGCGGTGGTCAAT 1362
Qy 1321 CAGATTAATGATAACCAATGAAACAGATCGCGGAGCGCTTCTTCCAGCCCGCTCGCGC 1380
Db 1363 CAGATTAATGACACCAATGAGAGGATCGGCGGAGCTGCGCTGCGCTGCGGCTGCTCGG 1422
Qy 1381 CTTTCTCTGCTTCGAGCTAATGATGCTTTGGCTCTCTCTCAGCGCTGCGGAGTAT 1440
Db 1423 CTTTCTCTGCTTCGAGCAATGATGTTGCTTTGGCTGTCCCTCACTGCGAGCGGAGTAT 1482

Qy 1441 GACCACTCCACTTATGGCTCTTGAATGCGCCAGTTTATGTTTCTGACTCTCTGACCTTG 1500
Db 1483 GACCACTCCACTTACGGGTCTGTAACCTGCGCGGTTTATATCTCGGACAGCGTGAATTG 1542
Qy 1501 GTTAATGTTGGAACGCGGCGGAGCGGCTTGGCCGCTGCTGATGAGCAAGTGCACA 1560
Db 1543 GTGAATGTTGCACTGCGGCGGAGCGGCTAGCCGATCGCTTGAAGTGCAGGCTCAC 1602
Qy 1561 CTTGAGGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTCTGCTGCCG 1620
Db 1603 CTCGACGCGCGCCCTCTCCGACTGTGAGCAATTTCCAGACATTTCTTGTGCTCCC 1662
Qy 1621 CTCGCGGTGAAGCTCTCTTTCTGGGAGGAGGACACAACTTAAAGCCGGGTACCTTATAAT 1680
Db 1663 CTTGTTGGAAGCTCTCTTTTGGGAGGCGCGCACAAACAAAGCAGGTTATCTTATAAT 1722
Qy 1681 TATAACACACTGTAGCAGCAAACTGTTGCGAGATGCGCGGCGGACCGGCTCGCT 1740
Db 1723 TATAATACTACTGTAGTGACCAATTTGATTTGAAATGCTGCGCGCATCGGCTCGCC 1782
Qy 1741 ATTTCACCTTACCACTAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1783 ATTTCACCTTACCACTAGGCTTGGGCGGCTTGGGCGGCTGCGCTGCGCATTTCTGGGCGCGGTT 1842
Qy 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGAGCTTACCTGCGCGCGC 1860
Db 1843 TTGCTTCCAGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1902
Qy 1861 CATACTTTTGTGATTTTCTGCGGAGTGGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTTC 1920
Db 1903 CACACATTTGATGACTTCTGCGCTGAAATCGCGGCTTTAGGCTTCCAGGCTTGGCTTTC 1962
Qy 1921 CAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 1963 CAGTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2022
Qy 1981 TAG 1983
Db 2023 TAG 2025

RESULT 11

US-09-851-410-10
; Sequence 10, Application US/09851410
; Publication No. US20030124510A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; Yarbough, Patrice O
; Bradley, Daniel W
; Krawczynski, Krzysztof Z
; Tam, Albert
; Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; No. US20030124510A1-A/No. US20030124510A1-B Hepatitis Viru
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,410
; FILING DATE: 07-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/128,275

6734	CTTCGTGGCAAGCTCTCCTTTTGGAGGCGCGGCACAAACAAAGCAGGTTATCCTTATAT	6793	QY	301	AGACCTACCAAGCTGGGGCGCGCGCTAAACCGCGGTGCTCGGGCCCATACACCCCG	360	
1681	TATAAACACCACTGCTAGCAGCAACTGCTTTGTTCGAGAAATGCCCGGGCACCGGGTCGCT	1740	Db	5417	CGACCTGCCACAGCGCGGGCTGGCGGCTGACGGCTGTGGCGCTGCCCATGACACCTCA	5476	
6794	TATAATACTACTGCTAGTGACCAAGATCTTGATTGAAAAATGCTCGCGGCCATCGGGTCGCC	6853	QY	361	CCAGTGCCTGATGTGCACTCCCGCGGCGCCATCTTGGCGCGGAGTATAACCTATCAACA	420	
1741	ATTTCACACTTACACCACTAGCCTGGGTGCTGTGTCGGTCTCCCGTCTCCATTTCTGCGGT	1800	Db	5477	CCCGTCCCGGACGTTTGATTCTTCGCGGTGCATTTCTACGCGCGCAGTATAAATTTGTCTACT	5536	
6854	ATTTCACACCTATACACACAGGCTTTGGGGCGCGTCCGGTCGCCATTTCTGCGCGCGGTT	6913	QY	421	TCCTCCCTTACCTCTTCGCTGGCCACCGGCACTAACTTGTTCTTTATGCGCGCCCTCTT	480	
1801	TTAGCCCCCACTCTGCGTAGCATCTGTTGAGGATACCTTGGACTACCTCGCCCGCGCC	1860	Db	5537	TCACCCCTTGACATCTCTGTGGCTCTTGGCATAATTTAGTCTCTGTATGACGCCCTT	5596	
6914	TTGGCTCCACGCTCCGCCCTGGCTCTGCTCGAGGATATCTTTGATTATCCGGGGCGGGCG	6973	QY	481	AGTCGCTTTTACCCCTTCAGGACGGCACCAATAACCATATATATATGAGCCACGGAAGCTTCT	540	
1861	CATACCTTTTGATGATTTCTGCCAGAGTGCGGCCCTTGGCCTTCAGGGCTGCGCTTTC	1920	Db	5597	AATCGCCTCTGCGCTGCAGACGGTACTAATACTACATTTATGSCCAGAGAGGCTCC	5656	
6974	CACACATTTGATGACTTCTGCCCTGAATGCGGCGCTTTAGGCCCTCCAGGGTTGTGCTTTC	7033	QY	541	AATTATGCCCACTACCGGGTTCGGGTGCGCAAAATCCGTTACGCCCGCTGGTCCCAAT	600	
1921	CAGTCTACTGTGCTGAGCTTCAGCGCTTAAAGATGAAGGTGGGTAAAACTCGGAGTTG	1980	Db	5657	AATTATGCACAGTACCGGGTTCGGCGCTACTATCCGTTACCGGCCCTTAGTGCCTAAT	5716	
7034	CAGTCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGGTGGGTAAAACTCGGAGTTG	7093	QY	601	GCTGTGCGGGTTACGCCATCTCCATCTCATTTCTTGTGGCCAAACCAACACACCCCTACA	660	
1981	TAG 1983		Db	5717	GCAGTTGGAGGCTATGCTATATCCATTTCTTCTTGGCTCAACCAACACACCCCTACA	5776	
7094	TAG 7096		QY	661	TCCGTTGATATGAATTAATAACCTCGACGGATGTTGCTATTATTTAGTCCAGGCCGCAFA	720	
RESULT 12			Db	5777	TCTGTTGACATGAATTCATTTACTTCCACTGATGTCAGGATTTCTTGTCAACCTGGCATA	5836	
US-10-239-090A-49			QY	721	GCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAAACCAAGGCTGGCGCTCC	780	
Sequence 49, Application US/10239090A			Db	5837	GCATCTGAATTTGGTTCATCCAAAGCAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG	5896	
Publication No. US20040101820A1			QY	781	GTCGAGACCTCTGGGTGGCTGAGGAGGAGGCTACCTCTGCTCTGTGTTATGCTTTGCATA	840	
GENERAL INFORMATION:			Db	5897	GTTGAGACATCTGGTGTGCTGAGGAGGAGGAGCACCCTCCGGTCTTGTGATGTTATGCATA	5956	
APPLICANT: KABUSHIKI KALSHA TOSHIBA			QY	841	CATGGCTCACTCGTAAATTCCTATATAATAACACCTTATACCGGTGCCCTCGGGCTGTG	900	
TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co			Db	5957	CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG	6016	
TITLE OF INVENTION: from Japanese, chip including the same, kit including the same,			QY	901	GACTTTGCCCTTGAGCTTGTGGAACCTTACCCCGGTAAACCAACCAATACCGGCTCG	960	
TITLE OF INVENTION: detecting hepatitis E virus using the same			Db	6017	GACTTTGCCCTTAGAGCTTGAGTTTCGCAATCTCACCACCTGTAAACCAATAACAGTGTG	6076	
FILE REFERENCE: 02S0741P			QY	961	TCCGTTATTTCACAGCACTGCTCGCCACCGCTTCGTCGCGGTGCGGACGGGACTCGCGAG	1020	
CURRENT APPLICATION NUMBER: US/10/239,090A			Db	6077	TCCGTTATTTCACAGCACTGCTCGTCGCTCACTCC---GCCCGAGGGCGCGACGCGGAG	6133	
PRIOR FILING DATE: 2003-07-24			QY	1021	CTCACCACCAACGGCTGTCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGTT	1080	
PRIOR FILING DATE: 2001-06-25			Db	6134	CTGACCACTACGAGCCACCGGTTTCATGAAAGATCTCCACTTACCGGCTTAATGGG	6193	
NUMBER OF SEQ ID NOS: 57			QY	1081	GTCGGTGAGATCGGCGCGGGATAGCCCTCACCTGTTCAACCTTGTTCGTCGACACTCTGCTT	1140	
SEQ ID NO 49			Db	6194	GTAAGTGAAGTTCGCGCGCGGATAGCTCTAACTTAACTTAACTTAACTTAACTTAACTT	6253	
LENGTH: 7180			QY	1141	GGCGGCTCGCGCAGAGAAATGATTTCGTGCGGTGCGGCGAGCTGTTCTACTTCCCGTCCC	1200	
TYPE: DNA			Db	6254	GGCGGCTCGCGCAGAGAAATTAATTTGTCGCGGTGCGGCGGCACTGTTTATTCGCGCGCG	6313	
ORGANISM: Hepatitis E Virus Mexico (M74506)			QY	1201	GTTGTCTCAGCCAAATGGCGAGCGGACTGTTAAAGTTGTATATCATCTGTAGAGAATGCTCAG	1260	
US-10-239-090A-49			Db	6314	GTTGTCTCAGCCAAATGGCGAGCCAAACGTTGAAAGCTCTATACATCAGTGGAGAATGCTCAG	6373	
Query Match			QY	1261	CAGATAAGGGTATTGCAATCCCGCATGACATGACCTCGGAGAACTCTCGTGTGGTATT	1320	
Best Local Similarity			Db	6374	CAGATAAGGGTATTGCTATCCCCCAGATATCGATCTTGTGATTTCGCGTGTGCTCATT	6433	
Matches 1606; Conservative			QY	1321	CAGGATTATGATAACCAACATGAACAAAGATCGGCGAGCGGCTTCTCCAGCCCAATCGCGC	1380	
1	ATCGGCTCGGCTATTTGTTGCTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA		60	Db	6434	CAGGATTATGACAAACGAGCATGACGAGATCGGCCCAACCCCGTTCGCTGCGCCATCTCGG	6493
5117			5176	QY	1381	CCTTTCTCTGTCTTCGAGCTAAATGATGTGCTTTGGTCTCTCTCTCACCGCTGCCAGTAT	1440
61	CCGCGCGGTCAGCGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG		120				
5177			5236				
121	TTCTGGGGTGACCGGGTGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC		180				
5237			5296				
181	CCCTTTCGCCCGGATGTCACCGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG		240				
5297			5356				
241	CGACCACTCGGCTCGGCTTGGCGTGACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG		300				
5357			5416				

Db 6313 GGTGGTCTGCCACGAAATTGATTTCTGGCTGGGGCCAGTGTCTTACTCCGCCCT 6372
Qy 1201 GTTGTCTCAGCAATGGCGAGCCGACTGTAAAGTTGTATACATCTCTAGAGAATGCTCAG 1260
Db 6373 GTCTCTCAGCAATGGCGAATGCGACTGTAAAGTTATATACATCTCTGTGAGAAATGCACAG 6432
Qy 1261 CAGGATAAGGGTATTCGAATCCGCGANGACATTTGACCTCGAGAAATCTCGTGTGGTATTT 1320
Db 6433 CAGGATAAGGGTATTCACCATCCACACGATATAGACCTCGGTGATTTCTGTGTGGTCAATC 6492
Qy 1321 CAGGATATGATAACCAATGACAGATCGCGACGCTTCTCAGCCCATCGCGC 1380
Db 6493 CAGGACTATGATAACAGACAGACGAGCAGGACCGACCTACCCCTCACCGCCCTCTCGT 6552
Qy 1381 CTTTCTCTGTCTCGAGCTAATGATGTGCTTTGGCTCTCTCTCAACCGCTGCCAGTAT 1440
Db 6553 CCAATCTCGGTTCTTCGCGCTAATGATGTCTGTGGCTCTCTCTTACCGCTGCCAGTAT 6612
Qy 1441 GACCACTGCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 6613 GACCACTGCACTATGGCTCATCTACCAACCCCATGTATGTCTCGGACACAGTCAATTT 6672
Qy 1501 GTTAATGTTGCGACCGCGCGAGCGTTGCGGCTCGCTCGATTGGACCAAGGTCAACA 1560
Db 6673 GTTAAAGTGGCCCACTGTGTCTCAGGCTGTGCTGCTCCCTTGATTTGGTCTAAAGTTACC 6732
Qy 1561 CTTGAGCGTCCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCTCGCG 1620
Db 6733 TTGACCGTTCGCCCTCTCACTATTTACAGCAGTATTTCTAAGACATTTCTACGTTCTCCCG 6792
Qy 1621 CTCGGCGGTAAAGCTCTTTCTGGGAGGAGGACCAACTAAAGCCGGTACCTTATAAT 1680
Db 6793 CTTGCGTGGAAAGTTGTCTTTTGGGAGGCTGGCACGACCAAGGCTGGTTATTCCTTTAAT 6852
Qy 1681 TATAACACCACTGCTAGCGACCACTGCTTTGTGAGAAATGCGCGGGCACCGGCTCGCT 1740
Db 6853 TATAATAACAATGCTAGTGATCAGATTTTAAATAGAGAACGAGCTGGCCATCTGTGCT 6912
Qy 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCGCTCTCCATTTCTGGGTTGCGGTT 1800
Db 6913 ATCTACTTACACTACCACTGCGCGCGCGGCCCACTTCTATCTCTGGGCTCGGTGTG 6972
Qy 1801 TTAGCCCCCACTCTGGCTAGCATTTGCTTGAGATACCTTGGACTACCTGCGCGGCC 1860
Db 6973 TTAGCCCGCACTCAGCCCTTGCTGTCTTTGAGGATACCGCGGATACCCCGCTCGTGCC 7032
Qy 1861 CATACTTTTGATGATTTCTGCCAGAGTTCGCGCCCTTGGCTTTCAGGGTTCGCGCTTTC 1920
Db 7033 CATACTTTTGATGACTTCTGCCCGGAATGTCGACCTTGGTTTTCAGGGCTGCGCTTTC 7092
Qy 1921 CAGTCTACTGTGCTGAGCTTTCAGCGCTTAAAGTAAAGTGGGTAAATCTCGGAGTTG 1980
Db 7093 CAGTCTACTATGCTGAGCTTTCAGCGCTTAAATGAAAGTAAAGTAAATCCCGGAGTTC 7152
Qy 1981 TA 1982
Db 7153 TA 7154

RESULT 14
US-10-239-090A-47
; Sequence 47, Application US/10239090A
; Publication No. US20040101820A1
; GENERAL INFORMATION:
; APPLICANT: KABUSHIKI KAISHA TOSHIBA
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co
; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same,
; TITLE OF INVENTION: detecting hepatitis E virus using the same
; FILE REFERENCE: 0250741P
; CURRENT APPLICATION NUMBER: US/10/239,090A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837

; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 57
; SEQ ID NO 47
; LENGTH: 7228
; TYPE: DNA
; ORGANISM: Hepatitis E Virus Japan JKK-Sap (AB074917)
US-10-239-090A-47

Query Match 68.0%; Score 1347.6; DB 17; Length 7228;
Best Local Similarity 80.1%; Pred. No. 0;
Matches 1584; Conservative 0; Mismatches 394; Indels 0; Gaps 0;

Qy 1 ATGCGCCCTCGCGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGTCGCGCGCA 60
Db 5165 ATGCGCTCTCGGCTTTTCTGTTTTTCTGCTCTGCTATGCTATGCTGCGCGCA 5224

Qy 61 CCGCCCGGTGACGCGCTCTGCGCGCGCTGTGGGCGGCGAGCGCGGTTCGCGCGTGT 120
Db 5225 CCGCGCGGTGACGCGCTCTGCGCGCGCTGTGGGCGGCGAGCGCGGTTCGCGCGTGT 5284

Qy 121 TTCTGGGGTGACCGGGTTGATTTCTCAGCCCTTCGCAATCCCTATATTTTCATCCAAACAAC 180
Db 5285 TTCTGGGGTGACCGGGTTGATTTCTCAGCCCTTCGCGCTTCCCTATATTTTCATCCAAACAAC 5344

Qy 181 CCCTTCGCCCCGATGTCAACCGCTGCGCGCGGGCTGGAGCTCTGTTGCGCCAAACCGCC 240
Db 5345 CCCTTCGCACTCTGACATTCAGCCGCGAGCGCGGGCTGGAGCTCGCCCTCGCGACGCGCC 5404

Qy 241 CGACCACTGCGCTTCGCGCTTGGCGTGACGAGGCCAGGCGCCCGCGCTGCTCAGCGTGT 300
Db 5405 CGTCCCACTCGGCTTCGCGCTTGGCGTGACCAATCCAGCGCGCCCGCCACTTCGCGCGCTGT 5464

Qy 301 AGACCTTACCACTGCGCGCGCGCTAAACCGCGCTCGCTCCGCGCCATGACACCCCG 360
Db 5465 CGATCTGCGCCACTGCGGGTTTCGCGCTGACTGTGTGCGCGCCCGCCAGATACTGT 5524

Qy 361 CCAGTGTCTGATGTGCTCCCGCGCGCGCAATCTTGGCGCGGAGTATACCTATCAACA 420
Db 5525 CCGTGTCCCGATGTGATTTCTCGCGCGCTATATTTAGCGCGCCAGTATAATTTATCAACA 5584

Qy 421 TCTCCCTTACTCTTCCGCTGGCGACCGGCACTAACTGCTTCTTTATGCGCGCCCTCT 480
Db 5585 TCCCGCGTAACTACTATTTGCGCACTGTGACTAACTTGTCTTATATGCTGCGCGCTG 5644

Qy 481 AGTTCGCTTTTACCCCTTCAGGACGGCAACCAATACCATATATATGCGCAAGAGTCT 540
Db 5645 AGCCCTTGTCTCGCTCCAGATGGAACCTAACTCACTCATATATGCGCACTGAAGATCA 5704

Qy 541 AATTATGCCAGTACCGGGTTCGCGCGCAATCCGTTACCGCGCGCTGCTGCTCCCAAT 600
Db 5705 AATTATGCCAGTACCGGTGTTGTCGCGCGCACCAATCCGCTACCGCTTGTGCGCAAC 5764

Qy 601 GCTGTCCGGGTTAGCGCATCTCCATCTCTATTCTGCGCACAGACCAACCCACCCGACG 660
Db 5765 GCTGTCCGGGATACGCTATATCTCTTCTTGGCTCTAGAACAACTACCAACCCGACCA 5824

Qy 661 TCCGTTGATATGAATTCAAATAACCTCGACCGGATGCTGTATTTTAGTCCAGCCCGGCATA 720
Db 5825 TCTGTGATATGAATTTCTATCACCTCCAGGATGCTCGGAATCTTGTTCAGCGCTGGCATC 5884

Qy 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTCACTATCTGTAAACCAAGGCTGGCGTCC 780
Db 5885 GCTTCAGAACTTGTGATCCCTTAGTGAGCGCTGCAATTTATGTAAACCAAGGCTGGCGTCC 5944

Qy 781 GTCGAGACCTCTGGGTTGGCTGAGGAGGCTACCTCTGCTCTGTTGTTATCTTTTGCATA 840
Db 5945 GTTGAGACCTCTGCTGTTTGGGAGGAGGAGGAGGAGCTCCGCGCTTGTCTATCTTTGCATC 6004

Qy 841 CATGCTCACTCTGTAATTCCTATATAATACCTTATACCGGTGCGCTCGGGCTGTG 900
Db 6005 CATGATCACTCTGTAATTTCTTACCAATACGCTTATATCTGGTGGCTTGGCTTGTCT 6064

Qy 901 GACTTTGCGCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTTAACCAACATACGCGGTC 960

Db	6967	CTAGCCCCACACTCGGCCCTTCGCCGCTCTAGAGGACACCGTTGATTACCCCGCCCGCGCT	7026
Qy	1861	CATACTTTTCATGATTTCTGCCCAGAGTGCCGCCCTTGCGCCTTCAGGGCTTCGCGTTTC	1920
Db	7027	CACACTTTTCATGATTTTTCGCCCGGAGTGCCGTACCTCGGTTTGCAGGGTTGTGCATTC	7086
Qy	1921	CAGTCTACTCTCGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTAAAACTCGGGAGTTG	1980
Db	7087	CAGTCTACTATCGCTGAGCTTCAGGCTCTTAAATGAAGGTAGTGTAACCCGGGAGTCT	7146
Qy	1981	TA	1982
Db	7147	TA	7148

Search completed: August 21, 2004, 22:41:57
Job time : 895.671 secs

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 08:42:47 ; Search time 7630.72 Seconds
(without alignments)
11246.543 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096

Perfect score: 1980

Sequence: 1 ATGGCGCCTAGGCTCTTTT.....GTAAACTCGGAGTTCTAG 1980

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	1980	100.0	1980	6	I73140	I73140 Sequence 2
2	1980	100.0	2055	6	AR002161	AR002161 Sequence
3	1980	100.0	2055	6	AR170405	AR170405 Sequence
4	1980	100.0	2055	6	AR232466	AR232466 Sequence
5	1980	100.0	2100	6	AR002159	AR002159 Sequence
6	1980	100.0	2100	6	AR232464	AR232464 Sequence
7	1980	100.0	2100	6	AR150916	AR150916 Sequence
8	1980	100.0	7171	6	AR207634	AR207634 Sequence
9	1980	100.0	7180	6	BD189811	BD189811 Polynucle
10	1980	100.0	7180	14	HPENSSP	M74506 Hepatitis E
11	1644	83.0	1644	6	AR170407	AR170407 Sequence
12	1378.2	69.6	1983	6	EJ7107	EJ7107 Hepatitis E
13	1378.2	69.6	7138	14	HPEA	M00581 Hepatitis E
14	1378.2	69.6	7168	6	AR139826	AR139826 Sequence
15	1378.2	69.6	7168	6	AR167470	AR167470 Sequence
16	1378.2	69.6	7168	6	AR234194	AR234194 Sequence
17	1378.2	69.6	7168	6	BD084498	BD084498 Recombina
18	1378.2	69.6	7194	6	EJ7109	EJ7109 Hepatitis E
19	1378.2	69.6	7204	14	AF444002	AF444002 Hepatitis
20	1376.6	69.5	2173	14	AF065061	AF065061 Hepatitis
21	1376.6	69.5	7204	14	AF444003	AF444003 Hepatitis
22	1376.6	69.5	7212	14	AY230202	AY230202 Hepatitis
23	1373.4	69.4	7193	14	HPEGENA	L25547 Hepatitis E
24	1373.4	69.4	7221	14	HPEORFS	L25595 Hepatitis E
25	1371.8	69.3	1984	6	AR278906	AR278906 Sequence
26	1371.8	69.3	7207	14	HPECG	D11092 Hepatitis E
27	1368.6	69.1	1983	6	I73139	I73139 Sequence 1
28	1368.6	69.1	1983	14	AF124407	AF124407 Hepatitis
29	1368.6	69.1	2049	6	AR002160	AR002160 Sequence
30	1368.6	69.1	2049	6	AR170404	AR170404 Sequence
31	1368.6	69.1	2049	6	AR232465	AR232465 Sequence
32	1368.6	69.1	2094	6	AR002158	AR002158 Sequence
33	1368.6	69.1	2094	6	AR232463	AR232463 Sequence
34	1368.6	69.1	7195	6	AR150915	AR150915 Sequence
35	1368.6	69.1	7195	6	AR207633	AR207633 Sequence
36	1368.6	69.1	7207	14	HPEVSP	M73218 Hepatitis E
37	1367	69.0	7170	14	AY204877	AY204877 Hepatitis
38	1365.4	69.0	7143	14	AF185822	AF185822 Hepatitis
39	1365.4	69.0	7194	14	HEVORFS	X99441 Hepatitis E
40	1365.4	69.0	7194	14	HPEHEV	D10330 Hepatitis E
41	1365.4	69.0	7194	14	HPEUIGH	D11093 Hepatitis E
42	1363.8	68.9	7176	14	HPEGENSA	L08816 Hepatitis E
43	1363.8	68.9	7200	14	HPECEGNOM	M94177 Hepatitis E
44	1362.2	68.8	7199	14	AF051830	AF051830 Hepatitis
45	1362.2	68.8	7202	14	HEVCG	X98292 Hepatitis E

ALIGNMENTS

RESULT 1						
LOCUS	I73140	1980 bp	DNA	linear	PAT 03-APR-1998	
DEFINITION	Sequence 2 from patent US 5686239.					
ACCESSION	I73140					
VERSION	I73140.1	GI:3009279				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1980)					
AUTHORS	Reyes,G.R., Tam,A.W. and Yarbough,P.O.					
TITLE	Hepatitis E virus peptides and methods					
JOURNAL	Patent: US 5686239-A 2 11-NOV-1997;					
FEATURES	Location/Qualifiers					


```
Db 1361 ACATTGTGAGTCTTGCCCTGAATGCGCGCTTTAGGCTCCAGGTTGTGCTTCCAG 1920
Qy 1921 TCACTGTGCTGAGTCCAGCGCCTTAAAGTTAAGTGGGTAAACCTCGGAGTTGTAG 1980
Db 1921 TCACTGTGCTGAGTCCAGCGCCTTAAAGTTAAGTGGGTAAACCTCGGAGTTGTAG 1980

RESULT 3
LOCUS AR170405 2055 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 2 from Patent US 6291641.
ACCESSION AR170405
VERSION AR170405.1 GI:17908364
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2055)
AUTHORS Fuerst,T.R., McAtee,C.Patrick., Yarbough,P.O. and Zhang,Y.-F.
TITLE Hepatitis E virus antigens and uses therefor
JOURNAL Patent: US 6291641-A 2/18-SEP-2001;
FEATURES
    location/Qualifiers
        1..2055
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 1980; DB 6; Length 2055;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGCCCTAGAGCTCTTTGCTGTGTTGTTCTCTGTTGTTCTGCTATGTTGCCCGGCCA 60
Db 1 ATGCGCCCTAGGCTCTTTGCTGTGTTGTTCTCTGTTGTTCTGCTATGTTGCCCGGCCA 60

Qy 61 CCGACCGGTGACGCGGTCTGCGCGCGCGTCTGCGCGCGCGCGCGCGGTACCGCGGTGTT 120
Db 61 CCGACCGGTGACGCGGTCTGCGCGCGCGTCTGCGCGCGCGCGCGGTACCGCGGTGTT 120

Qy 121 TTCTGGGGTACGCGGTGATTTCTAGCCCTTCGCGATCCCTATATTCATCAACCAAC 180
Db 121 TTCTGGGGTACGCGGTGATTTCTAGCCCTTCGCGATCCCTATATTCATCAACCAAC 180

Qy 181 CCTTTGCCCCAGAGCTTCGCGCTGCGTCTGACCTCTGACCTCGCTTCGCCAACCCAGCC 240
Db 181 CCTTTGCCCCAGAGCTTCGCGCTGCGTCTGACCTCTGACCTCGCTTCGCCAACCCAGCC 240

Qy 241 CCGCCACTTTGGTCCACTTTGGCGAGATCAGGCCCGCGCTGCGCTGCGCTCCCGTCCG 300
Db 241 CCGCCACTTTGGTCCACTTTGGCGAGATCAGGCCCGCGCTGCGCTGCGCTCCCGTCCG 300

Qy 301 CGACCTGCCACAGCGCGGCTGCGCGCTGACGGCTGCGCGCTGCGCGCTGCGCGCTGCG 360
Db 301 CGACCTGCCACAGCGCGGCTGCGCGCTGACGGCTGCGCGCTGCGCGCTGCGCGCTGCG 360

Qy 361 CCGCTGCCCGAGCTTGATTTCTCGCGTGCATTTCTACGCCCGCAGTAAATTTGTCTACT 420
Db 361 CCGCTGCCCGAGCTTGATTTCTCGCGTGCATTTCTACGCCCGCAGTAAATTTGTCTACT 420

Qy 421 TCACCCCTGACATCTCTGTGCGCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 480
Db 421 TCACCCCTGACATCTCTGTGCGCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 480

Qy 481 AATCCGCTCTGCGCTGCGAGCGGTACTATCTCAATTTAGTCTGATGATGCGCGCGCTCC 540
Db 481 AATCCGCTCTGCGCTGCGAGCGGTACTATCTCAATTTAGTCTGATGATGCGCGCGCTCC 540

Qy 541 AATTATGACAGTACCGGGTTCGCGGCTACTATCCGTTACCGGCGCTAGTGCCTAAT 600
Db 541 AATTATGACAGTACCGGGTTCGCGGCTACTATCCGTTACCGGCGCTAGTGCCTAAT 600

Qy 601 GCAGTTGGAGGCTATGCTATATCCATTCTTTCTGGCCTCAAAACCAACCCCTACA 660
```

```
Db 601 GCAGTTGGAGGCTATGCTATATCCATTCTTTCTGGCCTCAAAACCAACCCCTACA 660
Qy 661 TCTGTTGACATGAATTCATTACTTCCACTGATGTGAGGATTTCTGTTTCAACTGGCATA 720
Db 661 TCTGTTGACATGAATTCATTACTTCCACTGATGTGAGGATTTCTGTTTCAACTGGCATA 720

Qy 721 GCATCTGAATTTGCTCATCCCAAGCGAGCGCTTCACTACCGCAATCAAGTTGGCGCTCG 780
Db 721 GCATCTGAATTTGCTCATCCCAAGCGAGCGCTTCACTACCGCAATCAAGTTGGCGCTCG 780

Qy 781 GTTGAGACATCTGTTGTTGCTGAGGAGGAGCCACTCCGGTCTTTGTCTCATGTTATGCATA 840
Db 781 GTTGAGACATCTGTTGTTGCTGAGGAGGAGGAGCCACTCCGGTCTTTGTCTCATGTTATGCATA 840

Qy 841 CATGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACG 900
Db 841 CATGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACG 900

Qy 901 GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAAACACCAATACAGTGTG 960
Db 901 GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAAACACCAATACAGTGTG 960

Qy 961 TCCCGTTACTCCAGCACTGCTGCTCACTCCGCCGAGGGGCGAGCGGAGCTCGGAGCTG 1020
Db 961 TCCCGTTACTCCAGCACTGCTGCTCACTCCGCCGAGGGGCGAGCGGAGCTCGGAGCTG 1020

Qy 1021 ACCAACAATGAGCCAGCGGTTCATGAAGATCTCCACTTTACCGGCTTAAATGGGTA 1080
Db 1021 ACCAACAATGAGCCAGCGGTTCATGAAGATCTCCACTTTACCGGCTTAAATGGGTA 1080

Qy 1081 GGTGAAGTCGCGCGGGGATAGCTCTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 1140
Db 1081 GGTGAAGTCGCGCGGGGATAGCTCTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 1140

Qy 1141 GGGCTCCCGACAGAAATTAATTTGCTGCGTGGCGGGGCAACTGTTTAACTTAACTTAACTT 1200
Db 1141 GGGCTCCCGACAGAAATTAATTTGCTGCGTGGCGGGGCAACTGTTTAACTTAACTTAACTT 1200

Qy 1201 GTCTCAGCCAAATGGCGGCAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAGCAG 1260
Db 1201 GTCTCAGCCAAATGGCGGCAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAGCAG 1260

Qy 1261 GATAAGGTTGTTGCTATCCCCACGATATCGATCTTTGGTGAATTCGCTGTGTTCAATCAG 1320
Db 1261 GATAAGGTTGTTGCTATCCCCACGATATCGATCTTTGGTGAATTCGCTGTGTTCAATCAG 1320

Qy 1321 GATTATGACAAACAGCATGAGCAGGATCGGCCACCCCGTCCGCTGCGCTGCGCTGCGCT 1380
Db 1321 GATTATGACAAACAGCATGAGCAGGATCGGCCACCCCGTCCGCTGCGCTGCGCTGCGCT 1380

Qy 1381 TTTTCTGTTCTCCGAGCAAAATGATGTTACTTTGGCTGCTCCCTCACTGCGCGGATGATGAC 1440
Db 1381 TTTTCTGTTCTCCGAGCAAAATGATGTTACTTTGGCTGCTCCCTCACTGCGCGGATGATGAC 1440

Qy 1441 CAGTCCACTTACGGTTCGTTAACTGCGCCCGTTTATATCTCGGACAGCGTGAATTTGTTG 1500
Db 1441 CAGTCCACTTACGGTTCGTTAACTGCGCCCGTTTATATCTCGGACAGCGTGAATTTGTTG 1500

Qy 1501 AATGTTGGACCTGCGCGCGAGCGCTAGCCGATCGTTGACTGTTGCTGCTGCTGCTGCTG 1560
Db 1501 AATGTTGGACCTGCGCGCGAGCGCTAGCCGATCGTTGACTGTTGCTGCTGCTGCTGCTG 1560

Qy 1561 GACGGCGCGCCCTCCGACTGTTGAGCAATATTCAGAGCAATCTTTGTGCTGCTGCTGCT 1620
Db 1561 GACGGCGCGCCCTCCGACTGTTGAGCAATATTCAGAGCAATCTTTGTGCTGCTGCTGCT 1620

Qy 1621 CGTGGCAAGCTCTCCTTTTGGAGGCGCGCAACAAAGAGGTTATCCTTATATATAT 1680
Db 1621 CGTGGCAAGCTCTCCTTTTGGAGGCGCGCAACAAAGAGGTTATCCTTATATATAT 1680

Qy 1681 AATACTAGTCTAGTGAACAGATTTGATTTGAATGCTGCTGCGGCGATCGGTCGCATT 1740
```


Db	1501	AATGTTGCACTGGCGCGCAGGCGGTAGCCGATCGCTTGACTGGTCCAAAGTCACCCCTC	1560
Qy	1561	GACGGCGGCCCTCCCGACTGTTGAGCAATATTCCAAGACATTCTTTGTGCTCCCCCTT	1620
Db	1561	GACGGCGGCCCTCCCGACTGTTGAGCAATATTCCAAGACATTCTTTGTGCTCCCCCTT	1620
Qy	1621	CGTGGCAAGCTCTCTTTTGGAGCGCGGCACAAACAAAGCAGGTTATCTTATAATTAT	1680
Db	1621	CGTGGCAAGCTCTCTTTTGGAGCGCGGCACAAACAAAGCAGGTTATCTCTTATAATTAT	1680
Qy	1681	AATACTACTGCTAGTGACAGATTCTGATTGAAATGCTGCCGCCATCGGTCGCCATT	1740
Db	1681	AATACTACTGCTAGTGACAGATTCTGATTGAAATGCTGCCGCCATCGGTCGCCATT	1740
Qy	1741	TCAACCTATACACAGGCTTGGGCGCGTCCGCTCGCCATTCTGCGCGCGGTTTTG	1800
Db	1741	TCAACCTATACACAGGCTTGGGCGCGTCCGCTCGCCATTCTGCGCGCGGTTTTG	1800
Qy	1801	GCTCCACGCTCCGCGCTGCTGCTGGAGGATCTTTTGATTATCCGGGCGGCGGCAC	1860
Db	1801	GCTCCACGCTCCGCGCTGCTGCTGGAGGATCTTTTGATTATCCGGGCGGCGGCAC	1860
Qy	1861	ACATTTGATGACTTCTCCCTGAAATCCGCGCTTTAGGCTCCAGGGTTGTCTTCCAG	1920
Db	1861	ACATTTGATGACTTCTCCCTGAAATCCGCGCTTTAGGCTCCAGGGTTGTCTTCCAG	1920
Qy	1921	TCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTTGTAG	1980
Db	1921	TCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTTGTAG	1980
RESULT 5			
AR002159			
LOCUS	AR002159	2100 bp	DNA linear PAT 04-DEC-1998
DEFINITION	Sequence 2 from patent US 5741490.		
ACCESSION	AR002159		
VERSION	AR002159.1 GI:3963713		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2100)		
AUTHORS	Reyes, G.R., Bradley, D.W., Twu, J.-S., Purdy, M.A., Tam, A.W., Krawczynski, K.Z. and Yarbough, P.D.		
TITLE	Hepatitis E virus vaccine and method		
JOURNAL	Patent: US 5741490-A 2 21-APR-1998;		
FEATURES	Location/Qualifiers		
source	1..2100		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match 100.0%; Score 1980; DB 6; Length 2100;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGCGCCCTAGGCTCTTTTGTCTGTTTCTCTTGTCTTGTCTATGTTGCCCGGCCA	60
Db	46	ATGCGCCCTAGGCTCTTTTGTCTTGTCTCTTGTCTTGTCTATGTTGCCCGGCCA	105
Qy	61	CCGACCGGTGACCGGTTGATTTCTAGCCCTTCGGCGCGCAGCGCGTACCGGGGTGT	120
Db	106	CCGACCGGTGACCGGTTGATTTCTAGCCCTTCGGCGCGCAGCGCGTACCGGGGTGT	165
Qy	121	TTCTGGGGTGACCGGGTTGATTTCTAGCCCTTCGCAATCCCTATATTCATCAACCAAC	180
Db	166	TTCTGGGGTGACCGGGTTGATTTCTAGCCCTTCGCAATCCCTATATTCATCAACCAAC	225
Qy	181	CCCTTTGCCCGCCAGACGTTGCGCTCGGCTCGGCTGAGACTCGCTTCGCCAACCGAGCC	240
Db	226	CCCTTTGCCCGCCAGACGTTGCGCTCGGCTCGGCTGAGACTCGCTTCGCCAACCGAGCC	285
Qy	241	CGGCCACTTGGCTCCACTTGGCGAGATCAGGCCCGAGCGCCCTCCGCTCGCTCCCGTCGC	300

Db	286	CGGCGACTTGGCTCCACTTTGGCGAGATCAGGCCCGCCCTCGCTCGCTCCCGTCCG	345
QY	301	CGACCTGCCACAGCGCGGCTGCGGCGTGAAGGCTGCGGCGCTGCCCATGACACCTCA	360
Db	346	CGACCTGCCACAGCGCGGCTGCGGCGTGAAGGCTGCGGCGCTGCCCATGACACCTCA	405
QY	361	CCGCTCCCGAGCTTGAATTTCTCGCGTGCAATTTCTACGCCGCGAGTATATTTGCTACT	420
Db	406	CCCGTCCCGAGCTTGAATTTCTCGCGTGCAATTTCTACGCCGCGAGTATATTTGCTACT	465
QY	421	TCACCCCTGACATTCCTCTGTGGCTCTGGCACTAATTTAGTCTCTGTATCAGCCCCCTT	480
Db	466	TCACCCCTGACATTCCTCTGTGGCTCTGGCACTAATTTAGTCTCTGTATCAGCCCCCTT	525
QY	481	AATCGGCTCTCGGCTGAGAGCGGTACTATACTACATATTGCGCCACAGAGGCTCC	540
Db	526	AATCGGCTCTCGGCTGAGAGCGGTACTATACTACATATTGCGCCACAGAGGCTCC	585
QY	541	AATTTATGACAGTACCGGCTTCCCGCGCTACTATCCGTTACGGCCCTAGTGCCTAAT	600
Db	586	AATTTATGACAGTACCGGCTTCCCGCGCTACTATCCGTTACGGCCCTAGTGCCTAAT	645
QY	601	GCAGTTGGAGGCTATGCTATATCCATTTCTTTTGGCTTCAAAACAAACCAACCCCTACA	660
Db	646	GCAGTTGGAGGCTATGCTATATCCATTTCTTTTGGCTTCAAAACCAACCAACCCCTACA	705
QY	661	TCGTTGACATGAATTCATTTACTTCCACTGATGTCAGGATTTCTGTTCAACCTGGCATA	720
Db	706	TCGTTGACATGAATTTCCATTTACTTCCACTGATGTCAGGATTTCTGTTCAACCTGGCATA	765
QY	721	GCATCTGAAATGCTCATCCAAAGCGGCTTCACTACCGCAATCAAGGTTGGCGCTCG	780
Db	766	GCATCTGAAATGCTCATCCAAAGCGGCTTCACTACCGCAATCAAGGTTGGCGCTCG	825
QY	781	GTTGAGACATCTGGTGTGCTGAGGAGGAGCACTCCGGTCTTGTCTCATGTTATGCAFA	840
Db	826	GTTGAGACATCTGGTGTGCTGAGGAGGAGCACTCCGGTCTTGTCTCATGTTATGCAFA	885
QY	841	CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTCGCCCTTGGCTTAC	900
Db	886	CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTCGCCCTTGGCTTAC	945
QY	901	GACTTTGCCCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAAACCAATACACCTGTG	960
Db	946	GACTTTGCCCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAAACCAATACACCTGTG	1005
QY	961	TCGCGTTACTCCAGCATCTGCTGTCATCTCGCGCGAGGGCGGACGCGGACTGCGGAGCTG	1020
Db	1006	TCGCGTTACTCCAGCATCTGCTGTCATCTCGCGCGAGGGCGGACGCGGACTGCGGAGCTG	1065
QY	1021	ACCACAACTGACGCCACCAAGGTTTCATGAAGATCTCCACTTTTACCGGCTTAAATGGGTA	1080
Db	1066	ACCACAACTGACGCCACCAAGGTTTCATGAAGATCTCCACTTTTACCGGCTTAAATGGGTA	1125
QY	1081	GGTGAAGTCCGCGCGGGATAGCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAAT	1140
Db	1126	GGTGAAGTCCGCGCGGGATAGCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAAT	1185
QY	1141	GGGCTCCCGACAGAAATTAATTTCTGCTGGCTGGCGGCAACTGTTTATTTCCGCGCGGTT	1200
Db	1186	GGGCTCCCGACAGAAATTAATTTCTGCTGGCTGGCGGCAACTGTTTATTTCCGCGCGGTT	1245
QY	1201	GTCCTAGCCAAATGGCGAGCCACCGTGAAGCTCTATACATCAGTGGAGAACTCTCAGCAG	1260
Db	1246	GTCCTAGCCAAATGGCGAGCCACCGTGAAGCTCTATACATCAGTGGAGAACTCTCAGCAG	1305
QY	1261	GATAAGGCTGTTGCTATCCCCCAGCATATCGATCTTTGTTGATTCGGTGTGTCATTCAG	1320
Db	1306	GATAAGGCTGTTGCTATCCCCCAGCATATCGATCTTTGTTGATTCGGTGTGTCATTCAG	1365
QY	1321	GATTATGACAAACAGCATGAGCAGGATCGGCCCAACCCCTCGCTCGGCACTTCGGCTT	1380

Db 1186 GGGCTCCGACAGAATTAATTTTCGTGGCTGGCGGCAACTGTTTTATTTCCCGCCGGTT 1245
QY 1201 GTCTCAGCCATGGGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATCTCAGCG 1260
Db 1246 GTCTCAGCCAAATGGGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATCTCAGCAG 1305
QY 1261 GATAAGGGTGTGCTATCCCCACAGATATCGATCTTGGTGATTGGGCTGTGGTCAATTCAG 1320
Db 1306 GATAAGGGTGTGCTATCCCCACAGATATCGATCTTGGTGATTGGGCTGTGGTCAATTCAG 1365
QY 1321 GATTATGACACAGCATGAGCAGATCGGCCCAACCCCGTCGCCCTCGCCATCTGGCCT 1380
Db 1366 GATTATGACACAGCATGAGCAGATCGGCCCAACCCCGTCGCCCTCGCCATCTGGCCT 1425
QY 1381 TTTTCTGTTCTCGAGCAATATGATGATCTTTGGCTGTCCCTCACTGCAGCCGAGTATGAC 1440
Db 1426 TTTTCTGTTCTCGAGCAATATGATGATCTTTGGCTGTCCCTCACTGCAGCCGAGTATGAC 1485
QY 1441 CAGTCCACTTACGGGTGCGTCAACTGGCCCGGTTTATATCTCGGACAGCGGTGACTTTGGTG 1500
Db 1486 CAGTCCACTTACGGGTGCGTCAACTGGCCCGGTTTATATCTCGGACAGCGGTGACTTTGGTG 1545
QY 1501 AATGTTGCACTGGGCGCAGCGCTGAGCCCGATCGCTTGAATGCTGCCAAAGTCACCCCTC 1560
Db 1546 AATGTTGCACTGGGCGCAGCGCTGAGCCCGATCGCTTGAATGCTGCCAAAGTCACCCCTC 1605
QY 1561 GAGGGCGGCCCTCCCGACTGTGAGCAATATTCAGACATTCCTTTGCTGCCCTT 1620
Db 1606 GAGGGCGGCCCTCCCGACTGTGAGCAATATTCAGACATTCCTTTGCTGCCCTT 1665
QY 1621 CGTGGCAAGCTCTCTTTTGGGAGCGCGCACAAAGAGCAGGTATCTTATAATAT 1680
Db 1666 CGTGGCAAGCTCTCTTTTGGGAGCGCGCACAAAGAGCAGGTATCTTATAATAT 1725
QY 1681 AATACTACTGCTAGTGACAGATTCGTGATGAAATGCTGCCGCCCATCGGTCGCCAT 1740
Db 1726 AATACTACTGCTAGTGACAGATTCGTGATGAAATGCTGCCGCCCATCGGTCGCCAT 1785
QY 1741 TCACCTATACCAACAGGCTTGGGCGGTCGGTCCGCTCGCATTTCTCGCGCGCGGTTTG 1800
Db 1786 TCACCTATACCAACAGGCTTGGGCGGTCGGTCCGCTCGCATTTCTCGCGCGCGGTTTG 1845
QY 1801 GCTCCACGCTCGGCCCTCGCTGCTGAGGATACATTTTGAATATCGGGCGGCGGCAC 1860
Db 1846 GCTCCACGCTCGGCCCTCGCTGCTGAGGATACATTTTGAATATCGGGCGGCGGCAC 1905
QY 1861 ACATTTGATGACTTCTGCTGCTGATCGCGCTTTAGGCTTCAGGCTTCAGGCTTCCAG 1920
Db 1906 ACATTTGATGACTTCTGCTGCTGATCGCGCTTTAGGCTTCAGGCTTCAGGCTTCCAG 1965
QY 1921 TCACCTGCTGCTGAGCTCCAGCGCCTTAAAGTTAAGTGGTAAAGTAACTCGGAGTTGTAG 1980
Db 1966 TCACCTGCTGCTGAGCTCCAGCGCCTTAAAGTTAAGTGGTAAAGTAACTCGGAGTTGTAG 2025

RESULT 7
LOCUS AR150916 7171 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 10 from patent US 6229005.
ACCESSION AR150916
VERSION AR150916.1 GI:15115507
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 7171)
Reyes,G.R., Yarbough,P.O., Bradley,D.W., Krawczynski,K.Z., Tam,A.W.
and Fry,K.E.
TITLE DNA sequences of enterically transmitted non-A/non-B hepatitis
viral agent
JOURNAL Patent: US 6229005-A 10 08-MAY-2001;
FEATURES Location/Qualifiers
1..7171 source

/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 1980; DB 6; Length 7171;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTAGGCGCTCTTTTGGTGTGTTTCTCTTGTCTTGTCTATGTTGCGCGCCA 60
Db 5117 ATGCGCCCTAGGCGCTCTTTTGGTGTGTTTCTCTTGTCTTGTCTATGTTGCGCGCCA 5176
QY 61 CGACCGGTCAGCGCTCTGCGCGCGCTCGTGGCGCGCAGCGCGGTACCGCGGTGT 120
Db 5177 CGACCGGTCAGCGCTCTGCGCGCGCTCGTGGCGCGCAGCGCGGTATCCGCGGTGT 5236
QY 121 TTCTGGGTCAGCGGTTGATTTCTAGCCCTTCGCAATCCCTATATTTATCCAAACAAC 180
Db 5237 TTCTGGGTCAGCGGTTGATTTCTAGCCCTTCGCAATCCCTATATTTATCCAAACAAC 5296
QY 181 CCTTTTGGCCCGACAGCTTGCCTGCTGGGTCTGGACCTCGCCTTCGCCAACCCAGCC 240
Db 5297 CCTTTTGGCCCGACAGCTTGCCTGCTGGGTCTGGACCTCGCCTTCGCCAACCCAGCC 5356
QY 241 CGGCCACTTTGGCTCCACTTTGGCGAGATCAGGCCCGCAGCGCCCTCGCTCCGCTCGC 300
Db 5357 CGGCCACTTTGGCTCCACTTTGGCGAGATCAGGCCCGCAGCGCCCTCGCTCCGCTCGC 5416
QY 301 CGACCTGCCACAGCGCGGCTCGGGCTGACGGCTGTGGCGCTGCGCGCTGCCATGACACTCA 360
Db 5417 CGACCTGCCACAGCGCGGCTCGGGCTGACGGCTGTGGCGCTGCGCGCTGCCATGACACTCA 5476
QY 361 CCGCTCCCGGAGCTTGATTTCTCGGTGCAATTTCTACGCGCGCAGTATATTTGTCTACT 420
Db 5477 CCGCTCCCGGAGCTTGATTTCTCGGTGCAATTTCTACGCGCGCAGTATATTTGTCTACT 5536
QY 421 TCACCCCTGACATCTCTGTGGCTCTGGCACTAATTTAGTCTGTATGAGCGCCCTT 480
Db 5537 TCACCCCTGACATCTCTGTGGCTCTGGCACTAATTTAGTCTGTATGAGCGCCCTT 5596
QY 481 AATCGCCTCTCGCGCTCGAGACGGTACTAATCTACATTCATATGCGCAGAGGCGCTCC 540
Db 5597 AATCGCCTCTCGCGCTCGAGACGGTACTAATCTACATTCATATGCGCAGAGGCGCTCC 5656
QY 541 AATTATGACAGTACCGGGTTCGCCGCTACTATCCGTTACCGGCCCTTAGTGCCTAAT 600
Db 5657 AATTATGACAGTACCGGGTTCGCCGCTACTATCCGTTACCGGCCCTTAGTGCCTAAT 5716
QY 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTTCTGGCCTCAAAACAACACCCCTACA 660
Db 5717 GCAGTTGGAGGCTATGCTATATCCATTTCTTTCTGGCCTCAAAACAACACCCCTACA 5776
QY 661 TCTGTTGACATGAATTCCTACTTCCACTGATGTCAGGATTTCTGTTCAACCTGGCATA 720
Db 5777 TCTGTTGACATGAATTCCTACTTCCACTGATGTCAGGATTTCTGTTCAACCTGGCATA 5836
QY 721 GCATCTGAATGCTCATCCCAAGCAGCGCTTCACTACCGCAATCAAGGTTTGGCGCTCG 780
Db 5837 GCATCTGAATGCTCATCCCAAGCAGCGCTTCACTACCGCAATCAAGGTTTGGCGCTCG 5896
QY 781 GTTGAGACATCTGGTGTGCTGAGGAGGAGCCACTCCCGGTCTTGTCAATGTTATGATA 840
Db 5897 GTTGAGACATCTGGTGTGCTGAGGAGGAGGAGCCACTCCCGGTCTTGTCAATGTTATGATA 5956
QY 841 CATGGCTCTCCAGTTAACTCCCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 900
Db 5957 CATGGCTCTCCAGTTAACTCCCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 6016
QY 901 GACTTTGCGCTTAGAGCTTGAGTTTCGCAATCTCACCCCTGTAAACAACCAATACACGTTG 960
Db 6017 GACTTTGCGCTTAGAGCTTGAGTTTCGCAATCTCACCCCTGTAAACAACCAATACACGTTG 6076
QY 961 TCCCGTTACTCCAGCACTGCTGCTCACTCGCGCCGAGGGGCCAGCGGACTCGGAGCTG 1020

Db 5357 CGGCCACTTTGGCTCCACTTTGGCGAGATCAGGCCCGAGCGCCCTCCGCTCCGCTCGC 5416
QY 301 CGACCTGCCACAGCGGGGCTCGCGGCTGACGGCTGTCGGCGCTGCCCATGACACCTCA 360
Db 5417 CGACCTGCCACAGCGGGGCTCGCGGCTGACGGCTGTCGGCGCTGCCCATGACACCTCA 5476
QY 361 CCGCTCCCGGAGCTTTGATTTCCGCGGTGCAATTTCTACGCCGCGAGTATAAATTTGCTACT 420
Db 5477 CCGCTCCCGGAGCTTTGATTTCCGCGGTGCAATTTCTACGCCGCGAGTATAAATTTGCTACT 5536
QY 421 TCACCCCTGACATCCTCTGCTGCGCTCTGSCACTAAATTTAGTCTGATGACGCCCCCTT 480
Db 5537 TCACCCCTGACATCCTCTGCTGCGCTCTGSCACTAAATTTAGTCTGATGACGCCCCCTT 5596
QY 481 AATCCGCTCTCCGCTGCGAGGAGCTACTAATCTCTACATTTAGSCACAGAGGCTCC 540
Db 5597 AATCCGCTCTCCGCTGCGAGGAGCTACTAATCTCTACATTTAGSCACAGAGGCTCC 5656
QY 541 AATTATGACAGTACCGGTTGCCGCGCTACTATCCGTTACCGGCCCTTAGTGCTTAAT 600
Db 5657 AATTATGACAGTACCGGTTGCCGCGCTACTATCCGTTACCGGCCCTTAGTGCTTAAT 5716
QY 601 GCAGTTGAGGCTATGCTATATCAATTTCTTGGCCTCAAAACACACACCCCTACA 660
Db 5717 GCAGTTGAGGCTATGCTATATCAATTTCTTGGCCTCAAAACACACACCCCTACA 5776
QY 661 TCTGTTGACATGAATTCATTAATCTCCACTGATGTCAGGATTTCTTGTCAACCTGGCATA 720
Db 5777 TCTGTTGACATGAATTCATTAATCTCCACTGATGTCAGGATTTCTTGTCAACCTGGCATA 5836
QY 721 GCATCTGAATTTGGTCAATCCCAAGGAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 780
Db 5837 GCATCTGAATTTGGTCAATCCCAAGGAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 5896
QY 781 GTTGAGACATCGTGTTGCTGAGGAGGAGCCACTCGGCTTGTGATGATGACATA 840
Db 5897 GTTGAGACATCGTGTTGCTGAGGAGGAGCCACTCGGCTTGTGATGATGACATA 5956
QY 841 CATGCTCTCCAGTTAACTCCCTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 900
Db 5957 CATGCTCTCCAGTTAACTCCCTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 6016
QY 901 GACTTTGCTTAGAGCTTTAGTTTCGCAATCTCACCACTGTAAACCAATACAGTGTG 960
Db 6017 GACTTTGCTTAGAGCTTTAGTTTCGCAATCTCACCACTGTAAACCAATACAGTGTG 6076
QY 961 TCCGTTTACTCCAGACTGCTGTCACCTCCGCGCGAGGGCGGAGGACTGCGGAGCTG 1020
Db 6077 TCCGTTTACTCCAGACTGCTGTCACCTCCGCGCGAGGGCGGAGGACTGCGGAGCTG 6136
QY 1021 ACCACAACTGCAGCCACCAAGTTTCATGAAAGATCTCCACTTTACCGGCTTAAATGGGGTA 1080
Db 6137 ACCACAACTGCAGCCACCAAGTTTCATGAAAGATCTCCACTTTACCGGCTTAAATGGGGTA 6196
QY 1081 GGTGAAGTTCGCGCGGATAGCTCTAATCTAATTAACCTTGTGACACGCTCTCCTGGC 1140
Db 6197 GGTGAAGTTCGCGCGGATAGCTCTAATCTAATTAACCTTGTGACACGCTCTCCTGGC 6256
QY 1141 GGGCTCCGACAGATTAATTTGCTGCGCTGCGGCGAATGTTTTATCCCGCGGCTT 1200
Db 6257 GGGCTCCGACAGATTAATTTGCTGCGCTGCGGCGAATGTTTTATCCCGCGGCTT 6316
QY 1201 GTCTCAGCAATGGCGAGCCCAACCGTGAAGTCTATACATCAGTGGAGAAATGCTCAGAG 1260
Db 6317 GTCTCAGCAATGGCGAGCCCAACCGTGAAGTCTATACATCAGTGGAGAAATGCTCAGAG 6376
QY 1261 GATAAGGCTGTGCTATCCCCCAGATATCGATCTTGGTGTATTCGCGTGTGGTCAATTCAG 1320
Db 6377 GATAAGGCTGTGCTATCCCCCAGATATCGATCTTGGTGTATTCGCGTGTGGTCAATTCAG 6436
QY 1321 GATTATGACACACGATGACAGGATCGGCCACCCCGCTCGCTGCGCCATCTCGGCT 1380
Db 6437 GATTATGACACACGATGACAGGATCGGCCACCCCGCTCGCTGCGCCATCTCGGCT 6496

QY 1381 TTTTCTGTTCTCGGACCAATGATGATCTTTGGCTGTCCCTCACTGCACGCGATGATGAC 1440
Db 6497 TTTTCTGTTCTCGGACCAATGATGATCTTTGGCTGTCCCTCACTGCACGCGATGATGAC 6556
QY 1441 CAGTCCACTTACGGGTGCTCAACTGCGCCGGTTTTATCTCGGACAGCGTGACTTTGGTG 1500
Db 6557 CAGTCCACTTACGGGTGCTCAACTGCGCCGGTTTTATCTCGGACAGCGTGACTTTGGTG 6616
QY 1501 AATGTTGCACTGCGCGGAGCGCTAGCCGATCGCTTGACTGTTGCCAAAGTCAACCTTC 1560
Db 6617 AATGTTGCACTGCGCGGAGCGCTAGCCGATCGCTTGACTGTTGCCAAAGTCAACCTTC 6676
QY 1561 GACGGGCGCCCTCCGACTGTTGAGCAATATTTCCAGACATTTCTTGTGCTCCCTTT 1620
Db 6677 GACGGGCGCCCTCCGACTGTTGAGCAATATTTCCAGACATTTCTTGTGCTCCCTTT 6736
QY 1621 CTTGCAAGCTCTCTCTTTTGGAGCGCGGCACAAACAAAGCAGGTTATCTTTATTAATAT 1680
Db 6737 CTTGCAAGCTCTCTCTTTTGGAGCGCGGCACAAACAAAGCAGGTTATCTTTATTAATAT 6796
QY 1681 AATACTACTGCTAGTGACAGATTTCTGATGAAATGCTGCCGCCATCGGCTGCCATTT 1740
Db 6797 AATACTACTGCTAGTGACAGATTTCTGATGAAATGCTGCCGCCATCGGCTGCCATTT 6856
QY 1741 TCAACCTATACACAGGCTTGGGCGCGTCCGCTCGGCATTTCTCGGCCCGGGTTTG 1800
Db 6857 TCAACCTATACACAGGCTTGGGCGCGTCCGCTCGGCATTTCTCGGCCCGGGTTTG 6916
QY 1801 GTCACACGCTCGGCCCTGGCTCTGCTGGAGGATACCTTTGATTATCCGGGCGGGCGAC 1860
Db 6917 GTCACACGCTCGGCCCTGGCTCTGCTGGAGGATACCTTTGATTATCCGGGCGGGCGAC 6976
QY 1861 ACATTTGATGATCTTCTGCCCTGGAATGCGCGCTTTAGGCTCCAGGGTGTGCTTCCAG 1920
Db 6977 ACATTTGATGATCTTCTGCCCTGGAATGCGCGCTTTAGGCTCCAGGGTGTGCTTCCAG 7036
QY 1921 TCAACTGTCGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTTGAG 1980
Db 7037 TCAACTGTCGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTTGAG 7096

RESULT 10

HPENSSP
LOCUS Hepatitis E virus (mexican strain) structural proteins and a nonstructural polyprotein genes, complete cds. VRL 02-AUG-1993
DEFINITION
M74506
ACCESSION GI:330017
VERSION
KEYWORDS nonstructural protein; structural protein.
SOURCE Hepatitis E virus
ORGANISM Hepatitis E virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Hepatitis E-like viruses.
REFERENCE 1 (bases 1 to 7180)
AUTHORS Huang C.C., Nguyen D., Fernandez J., Yun K.Y., Fry K.E., Bradley D.W., Tam A.W. and Reyes G.R.
TITLE Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HEV)
JOURNAL Virology 191 (2), 550-558 (1992)
MEDLINE 93079857
PUBMED 1448913
COMMENT Original source text: Hepatitis E virus cDNA to genomic RNA.
FEATURES
source
1. 7180
/organism="Hepatitis E virus"
/mol_type="genomic RNA"
/db_xref="taxon:12461"
4. 5079
/note="putative"
/codon_start=1
/product="nonstructural polyprotein"
/protein_id="AA45730.1"
CDS


```
Db 6317 GTCTCAGCAATGGCGAGCCACCGGTGAAGCTCTATACATCATGTGAGAAATGCTCAGCAG 6376
QY 1261 GATAAGGGTGTGCTATATCCCCACCATATCGATCTTGTGATTCGGTGTGTCATTACAG 1320
Db 6377 GATAAGGGTGTGCTATATCCCCACCATATCGATCTTGTGATTCGGTGTGTCATTACAG 6436
QY 1321 GATTATGACAAACAGATGAGCAGGATCGGCCCAACCCCGCTGCTGCGCATCTCGGCCT 1380
Db 6437 GATTATGACAAACAGATGAGCAGGATCGGCCCAACCCCGCTGCTGCGCATCTCGGCCT 6496
QY 1381 TTTTCTGTCTCCGAGCAATGATGATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Db 6497 TTTTCTGTCTCCGAGCAATGATGATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6556
QY 1441 CAGTCCACTTACGGGTGCTCAACTGGCCGGTTTATATCTCGGACAGCGTACTTTGGTG 1500
Db 6557 CAGTCCACTTACGGGTGCTCAACTGGCCGGTTTATATCTCGGACAGCGTACTTTGGTG 6616
QY 1501 AATGTTGCACTGGCGGAGCCCGTAGCCCGATCGCTTGACTGGTCCAAAGTCAACCCTC 1560
Db 6617 AATGTTGCACTGGCGGAGCCCGTAGCCCGATCGCTTGACTGGTCCAAAGTCAACCCTC 6676
QY 1561 GACGGCGGCCCTCCGACCTGTTGAGCAATATTCAGACATCTCTTGTGTCGCCCTT 1620
Db 6677 GACGGCGGCCCTCCGACCTGTTGAGCAATATTCAGACATCTCTTGTGTCGCCCTT 6736
QY 1621 CGTGGCAAGCTCTCCTTTGGGAGCGCGCAACAAAGACAGGTTATCCCTATATATAT 1680
Db 6737 CGTGGCAAGCTCTCCTTTGGGAGCGCGCAACAAAGACAGGTTATCCCTATATATAT 6796
QY 1681 AATACTACTGCTAGTACAGATCTGATTGAAATGCTGCCGCAATCGGCTCGCCATT 1740
Db 6797 AATACTACTGCTAGTACAGATCTGATTGAAATGCTGCCGCAATCGGCTCGCCATT 6856
QY 1741 TCAACCTATACACAGGCTTGGGCGCGTCCGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1800
Db 6857 TCAACCTATACACAGGCTTGGGCGCGTCCGCTCGGCTCGGCTCGGCTCGGCTCGGCT 6916
QY 1801 GCTCACGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1860
Db 6917 GCTCACGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 6976
QY 1861 ACATTTGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 6977 ACATTTGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7036
QY 1921 TCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 7037 TCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7096
```

```
RESULT 11
ARL170407 ARL170407 1644 bp DNA linear PAT 17-DEC-2001
LOCUS
DEFINITION Sequence 4 from patent US 6291641.
ACCESSION ARL170407
VERSION ARL170407.1 GI:17908366
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1644)
AUTHORS Fuerst, T.R., McAtee, C. Patrick., Yarbough, P.O. and Zhang, Y.-F.
TITLE Hepatitis E virus antigens and uses therefor
JOURNAL Patent: US 6291641-A 4 18-SEP-2001;
FEATURES
Location/Qualifiers
source
1..1644
/mol_type="unassigned DNA"
ORIGIN
Query Match 83.0%; Score 1644; DB 6; Length 1644;
Best Local Similarity 100.0%; Pred. No. 0;
```

```
Matches 1644; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 GCTGTGGGCGCTGCGCATGACACCTCAACCGCTCCCGGACGTTGATTTCTCGCGGTGCAATT 393
Db 1 GCTGTGGGCGCTGCGCATGACACCTCAACCGCTCCCGGACGTTGATTTCTCGCGGTGCAATT 60
QY 394 CTACGCGGCGCAGTATAAATTTGTCTACTTTCACCCCTGACATCTCTGTGGCCCTCTGCGACT 453
Db 61 CTACGCGGCGCAGTATAAATTTGTCTACTTTCACCCCTGACATCTCTGTGGCCCTCTGCGACT 120
QY 454 AATTTAGTCTCTGTATGACGCCCCCTTAATCGCCCTCTGCCGCTGCGAGACGGTACTAAT 513
Db 121 AATTTAGTCTCTGTATGACGCCCCCTTAATCGCCCTCTGCCGCTGCGAGACGGTACTAAT 180
QY 514 ACTCATATATGGCCACAGAGGCTCCAAATATGACAGTACCGGTTGCCCGGCTACT 573
Db 181 ACTCATATATGGCCACAGAGGCTCCAAATATGACAGTACCGGTTGCCCGGCTACT 240
QY 574 ATCCGCTTACCGGCGCTTAGTGCTTAATGCAGTTGAGGCTATGCTATATCCCATTTCTTTC 633
Db 241 ATCCGCTTACCGGCGCTTAGTGCTTAATGCAGTTGAGGCTATGCTATATCCCATTTCTTTC 300
QY 634 TGGCCTCAAAACACAAACCCCTACATCTGTGATGAATTCATTTCCATTTCTCCACTGAT 693
Db 301 TGGCCTCAAAACACAAACCCCTACATCTGTGATGAATTCATTTCCATTTCTCCACTGAT 360
QY 694 GTCAGGATTTCTTGTTCACCTGGGCTAGCATCTGAATTTGGTTCATCCCAAGGAGCGCTT 753
Db 361 GTCAGGATTTCTTGTTCACCTGGGCTAGCATCTGAATTTGGTTCATCCCAAGGAGCGCTT 420
QY 754 CACTACCGCAATCAAGTTGGGCTCGGCTGAGCATCTGGTGTGCTGAGGAGGAGG 813
Db 421 CACTACCGCAATCAAGTTGGGCTCGGCTGAGCATCTGGTGTGCTGAGGAGGAGG 480
QY 814 ACCTCCGCTCTTGTATGATATGATATGATGCTCCAGTTAACTCTATACCAATACC 873
Db 481 ACCTCCGCTCTTGTATGATATGATATGATGCTCCAGTTAACTCTATACCAATACC 540
QY 874 CTTTATACCGTGGCTTGGCTTACTGCACTTGTGCTTAGAGCTTGTGCTTGTGCAATCTC 933
Db 541 CTTTATACCGTGGCTTGGCTTACTGCACTTGTGCTTAGAGCTTGTGCTTGTGCAATCTC 600
QY 934 ACCACCTGTAAACCAATACACGCTGCTCCGCTTACTCCAGCATGCTGCTGCTGCTGCTGCT 993
Db 601 ACCACCTGTAAACCAATACACGCTGCTCCGCTTACTCCAGCATGCTGCTGCTGCTGCTGCT 660
QY 994 CGAGGGCGGACGGGACTGCGGAGCTGACCAAACTGACGACCAACGAGTTTCAAGAGAT 1053
Db 661 CGAGGGCGGACGGGACTGCGGAGCTGACCAAACTGACGACCAACGAGTTTCAAGAGAT 720
QY 1054 CTCACATTTACCGGCTTAAATGAGGTAGGTGAGTGGCGCGGATAGCTCTTAACATTA 1113
Db 721 CTCACATTTACCGGCTTAAATGAGGTAGGTGAGTGGCGCGGATAGCTCTTAACATTA 780
QY 1114 CTTAAACCTTGTGACACGCTCTCGGCGGCTCCCGGACAGAAATTAATTTCTGCTGCTGCTGCT 1173
Db 781 CTTAAACCTTGTGACACGCTCTCGGCGGCTCCCGGACAGAAATTAATTTCTGCTGCTGCTGCT 840
QY 1174 GGGCAACTGTTTTATTCGCGCGCTTGTCTCAGCCAAATGGCGAGCCAAACGCTGAAGCTC 1233
Db 841 GGGCAACTGTTTTATTCGCGCGCTTGTCTCAGCCAAATGGCGAGCCAAACGCTGAAGCTC 900
QY 1234 TATACATCAGTGAGAAATGCTCAGCAGGATAAGGCTTGTGCTATPCCCCACGATATCGAT 1293
Db 901 TATACATCAGTGAGAAATGCTCAGCAGGATAAGGCTTGTGCTATPCCCCACGATATCGAT 960
QY 1294 CTTTGTGATTTGCGGTGTTGCTTCCAGGATATGACAAACGAGCATGAGCAGGATCGGCC 1353
Db 961 CTTTGTGATTTGCGGTGTTGCTTCCAGGATATGACAAACGAGCATGAGCAGGATCGGCC 1020
QY 1354 ACCCGCTGCGCTGCGCCATCTCGGCCCTTTTCTGCTTCTCGGACAAATGATGACTTTGG 1413
Db 1021 ACCCGCTGCGCTGCGCCATCTCGGCCCTTTTCTGCTTCTCGGACAAATGATGACTTTGG 1080
```

Qy	1414	CTGTCCCTCACTGCAGCCGAGTATGACACAGTCCACTTACGGGTCTCAACTGGCCCGGTT	1473
Db	1081	CTGTCCCTCACTGCAGCCGAGTATGACAGTCCACTTACGGGTCTCAACTGGCCCGGTT	1140
Qy	1474	TATATCTCGAGCAGCGTGACTTTTGGTGAATGTTCGACTTGGCGCGCAGCGCGTAGCCCGA	1533
Db	1141	TATATCTCGAGCAGCGTGACTTTTGGTGAATGTTCGACTTGGCGCGCAGCGCGTAGCCCGA	1200
Qy	1534	TCGCTTGACTGTTCACAACTACCTTCGACGGCGCGCCCTCCGACACGTTTGAGCAATAT	1593
Db	1201	TCGCTTGACTGTTCACAACTACCTTCGACGGCGCGCCCTCCGACACGTTTGAGCAATAT	1260
Qy	1594	TCCAAGACATTTCTTGTGCTCCGCTTCGTGGCAAGCTCTCCTTTTGGGAGCGCGGCACA	1653
Db	1261	TCCAAGACATTTCTTGTGCTCCGCTTCGTGGCAAGCTCTCCTTTTGGGAGCGCGGCACA	1320
Qy	1654	ACAAAAGAGGTTATCCTTATTAATAATACTACTGTAGTAGACAGATTTCTGATTGAA	1713
Db	1321	ACAAAAGAGGTTATCCTTATTAATAATACTACTGTAGTAGACAGATTTCTGATTGAA	1380
Qy	1714	AATGCTGCCGGCCATCGGTGCGCAATTCAACTATACACAGAGCTTGGGCGCGGTCCG	1773
Db	1381	AATGCTGCCGGCCATCGGTGCGCAATTCAACTATACACAGAGCTTGGGCGCGGTCCG	1440
Qy	1774	GTGCGCATTTCTCGCGCCGCGGTTTTGGTTCACAGCTCCGCGCTCTGCTGGAGGAT	1833
Db	1441	GTGCGCATTTCTCGCGCCGCGGTTTTGGTTCACAGCTCCGCGCTCTGCTGGAGGAT	1500
Qy	1834	ACTTTTGAATATCCGGGGCGGCGCACACATTTGATGACTTCTGCCCTGAAATGCGCGCT	1893
Db	1501	ACTTTTGAATATCCGGGGCGGCGCACACATTTGATGACTTCTGCCCTGAAATGCGCGCT	1560
Qy	1894	TTAGGCTCCAGGGTTGTCTTTCCAGTCAACTGTCTGAGCTTCAGGGCTTAAAGTT	1953
Db	1561	TTAGGCTCCAGGGTTGTCTTTCCAGTCAACTGTCTGAGCTTCAGGGCTTAAAGTT	1620
Qy	1954	AAGTGGGTAAACTCGGGAGTTG	1977
Db	1621	AAGTGGGTAAACTCGGGAGTTG	1644
RESULT 12			
LOCUS	E17107	1983 bp	DNA linear PAT 28-JUL-1999
DEFINITION	Hepatitis E virus ORF2.		
ACCESSION	E17107		
VERSION	E17107.1 GI:5711790		
KEYWORDS	JP 1998234383-A/3.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1983)		
AUTHORS	Takeda,N., Ri,T. and Miyamura,T.		
TITLE	HEPATITIS E VIRUS HOLLOW PARTICLE, GENE ENCODING THE SAME AND PREPARATION OF RECOMBINATION VECTOR CONTAINING THE GENE AND UTILIZATION OF THE RECOMBINATION VECTOR		
JOURNAL	Patent: JP 1998234383-A 3 08-SEP-1998;		
COMMENT	DENKA SEIKEN CO LTD, KOKURITSU KANSENSHIYU KENKYUSHO		
	OS Hepatitis E virus		
	PN JP 1998234383-A/3		
	PD 08-SEP-1998		
	PF 28-FEB-1997 JP 1997062445		
	PI TAKEDA NAOKAZU, RI TENSEI, MIYAMURA TATSUO		
	PC C12N15/09,C07K14/08,C12N5/10,C12P21/02;		
	CC strandedness: Double;		
	CC topology: Linear;		
	FH Key		
	FH Location/Qualifiers		
FT	source	1..1983	Hepatitis E virus' FT CDS
FT	/organism='Hepatitis E virus'		
FT	/product='structural protein'.		
FT	1..1983		

901	Db	GAC	TTTGGCCCTTGAGCTTGAGTTTTCGCAACCTTACCCCGCGGTAAACAACAATACGCGGGTC	960
961	Qy	TCC	CGTTACTCTACGACACTGCTCGTCACTCC---GCCCGAGGGCCGACGGGACTCGCGAG	1017
961	Db	TCC	CGTTATTCCAGACACTGCTCGCAACCGCCTTCGTGCGGTGCGGACGGGACTGCCGAG	1020
1018	Qy	CTG	ACCAACAACCTGCAGGCACACAGGTTTCATGAAGAATCTCCACTTTACCGGCTTTAATGGG	1077
1021	Db	CTC	ACCACACACGGCGCGTTACCGGCTTTTGAAGGACCTCTATTTTACTAGTACTAATGGT	1080
1078	Qy	GTA	GGTGAAGTGGCGCGGATAGCTCTAACATATCTTAACCTTGCTGACAGCTCCTC	1137
1081	Db	GTC	CGTGAGATCGGCGCGGATAGCCCTCAACCTGTTTAAACCTTGCTGACACTCTGCTT	1140
1138	Qy	GGC	GGGCTCCCCGACAGAAATTAATTTCTGTCGCTGGCGGCAACTGTGTTTATTTCCGCGCCG	1197
1141	Db	GGC	GGCTCGCGACAGAATTGATTTCTGTCGCTGGTGGCCAGCTGTTCTACTCCGTCCTC	1200
1198	Qy	GTT	GTCAGCAATGGGAGCCAAACCGTGAAGTCTTATACATCAGTGGAGAAATGCTCAG	1257
1201	Db	GTT	GTCCTGGCCAATGGCGAGCCGACTGTTAAAGTTGTATACATCTGTAGAGAATGCTCAG	1260
1258	Qy	CAG	ATAAGGGTGTGCTATCCCCCAAGATATCAATCTTGGTGAATCGCGTGGTGCTCATTT	1317
1261	Db	CAG	ATAAGGGTATTTGCAATCCCGCATGACATTCACCTCGAGAAATCTCGTGTGGTTATT	1320
1318	Qy	CAG	ATTATGACAAACAGCATGACAGATCGGCCACCCCGTCGCTGGCCCATCTCGG	1377
1321	Db	CAG	ATTATGACAAACAATGACAGATCGGCCGAGCCTTCTCAGCCCCATCGCCG	1380
1378	Qy	CTT	TTTTTCTGTTCCGAGCAAAATGATGTATCTTTGGCTGTCCCTCACTCAGCGCGAGTAT	1437
1381	Db	CTT	TCTCTGTCCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACGCTGCGCGAGTAT	1440
1438	Qy	GAC	AGTCCCACTTACGGGTGCTCAACTGGCCCGGTTTATATCTCGGACAGGTGACTTTG	1497
1441	Db	GAC	AGTCCCACTTATGGCTCTTGCACTGGCCCACTTTATGTTTCTGACTCTGTGACCTTG	1500
1498	Qy	GTG	AATGTTGCGACTGGCGCGAGCGCGTAGCCGATCGCTTGACTGCTCCGTCCTCAAGTCACC	1557
1501	Db	GTT	TAATGTGGACCGGCGGAGCGCTTGCCCGCTCGCTTCGANTGGACAAGGTGACA	1560
1558	Qy	CTG	ACGGGGGCCCTCCGACTGTTTGAGCAATATTTCCAAGACATTTCTTTGTGCTCCCC	1617
1561	Db	CTT	GAOGGTCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTGCTCTGCCG	1620
1618	Qy	CTT	CGTGGCAAGCTCTCTTTTGGAGGCGCGGACAAACAAAGAGAGTTATTCCTATAAAT	1677
1621	Db	CTC	CGCGGTAAAGCTCTCTTTCTGGAGGCGAGGCACAACTAAAGCCGGGTACCCCTATAAT	1680
1678	Qy	TAT	AATACTACTGCTAGTGACAGATCTCATTTGAAAACTGTCGCCGCATCGGGTCGCC	1737
1681	Db	TATA	CACACTGCTAGGACCAACTGCTTGTTCAGAAATCGCCCGGCGACCCGGGTGCT	1740
1738	Qy	ATT	TCAACTATACCAACAGGCTTGGGCGCGGTCCGGTCGCCATTTCTGGGCGCGGGTT	1797
1741	Db	ATT	TCCACTTACACCACTAGCCTGGGTCGTCTCCCGTCTCCATCTCTCGGCTGCTGTT	1800
1798	Qy	TTG	CTCCAGCTCCGCGCCTGGCTCTCTCGAGAGGATCTTTTGATTATTCGGGGCGGCG	1857
1801	Db	TTA	GCCCCCACTCTGCGCTAGCATTCCTTTGAGGATACCTTTGGACTACCCGTGCCCGCGCC	1860
1858	Qy	CAC	ATTTGATGACTTCTCCCTGAATCGCGGCTTTAGGCTCCAGGCTTGTCCTTTC	1917
1861	Db	CAT	ACTTTTGATGATTTCTGCCAGAGTGGCGGCTCTTGCGCTCCAGGCTCGGCTTTC	1920
1918	Qy	CAG	TCAACTGTCTGCTGAGCTCCAGCGCTTTAAAGTTAAAGTGGGTAAACTCTCGGAGTTG	1977
1921	Db	CAG	TCTACTGCTGAGCTTCAGCGCTTTAAGATGAAGTGGGTAAACTCTCGGAGTTTA	1980
1978	Qy	TAG	1980	
1981	Db	TAG	1983	

RESULT 13
HPEA
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
FEATURES
source
CDS
CDS
CDS

```

HPEA                                7138 bp ds-RNA                    linear   VRL 02-AUG-1993
Hepatitis E Virus, complete cds.
M80581
M80581.1 GI:329997

.
Hepatitis E virus
Hepatitis E virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Hepatitis
E-like viruses.
1 (bases 1 to 7138)
Tsarev,S.A., Emerson,S.U., Reyes,G.R., Tsareva,T.S., Legters,L.J.,
Malik,I.A., Icabal,M. and Purcell,R.H.
Characterization of a prototype strain of hepatitis E virus
Proc. Natl. Acad. Sci. U.S.A. 89 (2), 559-563 (1992)
92115700
1731327
Original source text: Hepatitis E virus (strain SAR-55) cDNA to
genomic RNA.

Location/Qualifiers
1..7138
/organism="Hepatitis E virus"
/mol_type="genomic RNA"
/strain="SAR-55"
/db_xref="taxon:12461"
1..5079
/note="ORF-1"
/codon_start=1
/protein_id="AAA45725.1"
/db_xref="GI:329998"
/translation="EAAQPIKAPGITTAAQALAAANSALANAVVVRPFLSHQOQLI
LINMQPQLVPRPEVFNHNP1QRVHNELEYCPARGSCLEIGHAPRSINDNPNRV
HRCFLRPAGRDVQRVRYTAPTRGPAANCRRSALRGLPAADRTYCFDGFSGCNFPAETGI
AIDYLDHMSPSDVAAGAMFHGMTRLYAALHLPPEVLLPPTGYTASLYLLIHDGRRVVV
TYEGDTAGYNHDVSNLRSWIRTKVTGDHPLVIERVAIGHGVHLLTAAPSPSPV
VYPRSTEVVYRSYIFGPGTSLPFTSCSTKTHAVPAHWIDRLMFGATLDQAF
CCSRULTLRGISYKVTGTVLSEWGNWASDALTAVTAAVLTICHQVLRTOALSK
GMRLEERHAQKFIIRLYSWLPEKSGRDY1PGQOLEFYAQCRRMLSAGHELDPRVLVF
DESACHCETAIRKAVSXFCEPMKWLQBCETCFELQPAEYVGQDHDNEAYEGSDVDP
AESALDSGSTVYSGPTGLQPLQYALDPAE1VARAGRATATVVKVSDQGRIDCETL
GNKTRTFSDVAGVLENGPERKSNLSFDASQSTAAQPSFLTYAASAGLEGVRYAAG
LDHRAVAPQGVSEPAQEVTAFCNALYFNREAOQLSLTGNFWFPEGLLGPFAFAP
PGHWESANPFCGESLTLYRTWSEVDVAPSPAQGLFTSEPSISRAATPTPAALPR
PPADPPTLSAPAGEPAPGATARAPA1THOTARRHLLFTYPDGSKVFAQSLPEST
CTVLNANVDNHRPGGLCHAFQYPSAFSDAASFVMDGAAAYTLTPRPIHAVAPD
YLRLEHPRLEAYRETCRSLGTAAYPILGTY1QVPIGFSFMDVGRACAGCRVTPGVVQYQF
ELAAKWFANRPTLTITEDVATANLAL1ELDSATDVGACAGCRVTPGVVQYQF
AGVPGSKRSRITQADVVVVVPTRELNANRRGFAPFPTHTAARTVQGRVVIDEA
PSLPHELLLHMQRATVHLLGDNPQI1PAIDEPHAGLVPAIRPDLAPTSMWHVTHRCF
ADVTII1AGAPMOTTISQSRRAH1VALTRHTEKCVI1DAAGLLREVGISDA1VHAQGT
YUETTT1IDAARGILISQSRRAH1VALTRHTEKCVI1DAAGLLREVGISDA1VHAQGT
AGGETGHORPSVI1PRGNDANVD1LAA1PPSCQ1SAPHQLABELGHRPAPVAAL1PPC
PELEGGLY1FQELTICDSV1TFLTD1VHCNMAAPQKRAVLSTLVGRYGRRTKLYN
ASHSDRSLRSLFAIPAGVQVVTTCLEYELVEAMVEKQDGSAVLELDLCSDRSVSRIT
FQKDCNKFTTGETTIAHGKVGQGI1SAWSKTFCALFQWFRATEKA1LALLPGCVFYGD
AFDQDVFSAAGAAASAAVFNDFSEFQSTNNPSLGLECA1MEECGMPQWLIR1LVLH
IRSAWILQAPKESLGRKFKHKSGBPGTLLWNTVNMNAVITHCYDFRDLQVAFKGDSD
IYLCSEYRQSGCAAVL1AGCGILKLVKDFRPTGLYAGVVAVPGLLGALPDVVRFAGRLTE
KNWGGPRRQSLRIASVDFLRKLTN1VQAQMVVDVSVYGVSGPLVHNL1IGMLQAVAD
GKAHFTESVXPVLDLTNSILCRVE"
5076..5447
/note="ORF-3"
/codon_start=1
/protein_id="AAA45726.1"
/db_xref="GI:329999"
/translation="MNNNSFAAPMGSRPCALGLFCCSSCFCLCPRHPRVSRSLAAVY
GGAAAPAVVSGVGTGLILSPSQSF1QFTSPSPMSPLRPGLLDVFANPPDHSAPLGV
TRPSAPPLVQVVDLPQLGPRR"
5117..7099
/note="ORF-2"

```



```
/codon_start=1
/protein_id="AAA5727.1"
/db_xref="GI:330000"
/translation="MRPRPILLLLMLFPLPMPAPPQPSGRRRRSGSGGGFWGD
RVDSQPAIPYIHPNTPFAPDVTAAGAGPRVRQPARPLGSAMRDQAPPAASRRRP
TTAGAAPLTAFAADHTPEVPDVSAGGAILRRQYNSPLTSSVATGNLVLVYAAPL
SPLLPLOGTNTNHWATBASNAOYRVARATYRPLVNPAGVVAISISFPWPOTTTT
PFSVDMNSTDVLILVQPGIASELVIPESELHYRNQMSRVETSGVAEEAATSLIV
MUCIHGSPVNSTYNTPIYTGALGLDFALEFRNLTPPNTNTRVRSYSSSTARHRRRG
ADGTAEALTTTAATRFMKDLYFTSNGVGEIGRIALTLFNLADTLGLLPTLELISAG
GQLFSRPFVVSANGBETVKLYTSVENAQDQKGIAPHIDLGEISRVIOYDNOHEQD
RTPSPAPSRPESLVRANDVLMLSLTAABYDOSYSSSTGVPVSDSVTLVNVATGAS
AVARSLDWTPLDRPLSTIOQYSKTPFVPLRGLKLSPEWAGTTKAGYPYNNYTAAS
DOLLVENAGHVAISTYTSLSGAGPVSI SAVAVLAPHSVLALLEDTMDYPARAHTFD
DFCPCEPLGLQGCAPQSVAELQRLMKVGTREL"
ORIGIN
Query Match          69.6%; Score 1378.2; DB 14; Length 7138;
Best local Similarity 81.3%; Pred. No. 0;
Matches 1612; Conservative 0; Mismatches 368; Indels 3; Gaps 1;
QY      1  ATGCGCCCTAGGCCTCTTTTGTGTTGTTCTCTGTTTCTGCTATGTTGCCCTATGTTGCCGGGCCA 60
Db      5117 ATGCGCCCTCGGCCTATTTTGTGTTGTTCTCTCATGTTTCTGCTATGTTGCCCTATGTTGCCGGGCCA 5176
QY      61  CCGACCGGTACGCGTCTGSCGCGCTGCTGGGCGCGAGCGGCGGTACCGGCGGTGTT 120
Db      5177 CCGCCCGGTACGCGTCTGSCGCGCTGCTGGGCGCGAGCGGCGGTTCCGGCGGTGTT 5236
QY      121 TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db      5237 TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 5296
QY      181 CCGTTTGGCCGACAGCTTGGCGTCTGCGGTCTGAGCTTGGACCTTGGCCACAGCC 240
Db      5297 CCGTTGCGCCGACAGCTTGGCGTCTGCGGTCTGAGCTTGGACCTTGGCCACAGCC 5356
QY      241 CGGCCACTTGGCTTCCACTTGGCGGAGATCAGGCCAGCGCCCTCGCTCGCTCCGCTCG 300
Db      5357 CGACCACTCGGCTCCGCTTGGCGGTGACGAGCCAGCGCCCGCGCTCGCTCACGCTGT 5416
QY      301 CGACCTGCCACAGCGGGGTGCGGGCTGACGGCTGTGGCGCTTCCCATGACACCTCA 360
Db      5417 AGACCTTACCACAGCTGGGGCGGCGCTAAACGCGGTGCTCGCGCCCATGACACCCCG 5476
QY      361 CCGGTCCCGGACGTTGATCTCCGGGTGCAATTTCTACGCGCGCAGTAAATTTGTCTACT 420
Db      5477 CCAGTGCCTGATGTTGACTCCCGCGCGCCATCTCGCGCGCAGTAAACCTTATCAACA 5536
QY      421 TCACCCCTGACATCTCTGTGGCTCTGGCACTAAATTTAGTCTCTGTATGAGCCGCCCTT 480
Db      5537 TCTCCCTCACTCTTCCGTGGGCAACGGGCAAAATTTGGTCTTTTACGCGCTCTCTTT 5596
QY      481 AATCCCGCTCTGCGCTGCGAGCGGTACTAATCTACATATATGCGCCACAGAGGCTCC 540
Db      5597 AGCCCGCTTCAACCCCTCCAGGACGCGACCAATCTCATATAATGCTACAGAACTTCT 5656
QY      541 AATTATGCACAGTACCGGGTTCGCGCTTACTATCCGTTTACGCGCCCTAGTGCCTAAT 600
Db      5657 AATTATGCCAGTACCGGGTTCGCGTGCACAAATTCGCTTACCGCCCGCTGGTCCCAAC 5716
QY      601 GCAGTTGGAGGCTATGCTATATCCATTTCTTCTGGCTTCAAAACCAACCCCTTACA 660
Db      5717 GCTGTTGGTGGCTACGCTATCTCAATTCGTTCTGGCCACAGACCAACCCCGAGC 5776
QY      661 TCTGTTGACATGAATTCCAATTAATTCACATGATGTCAGGATTCCTGTTCAACTGGGATA 720
Db      5777 TCGGTTGACATGAATTCATTAACCTGACGAGTGTCCGTATTTTATGTTCCAGCCCGGATA 5836
QY      721 GCATCTGAATTTGTCATCCCAAGGAGCGCTTCACTACCGCAATCAAGGTTGCGCGTCG 780
Db      5837 GCCTCCGAGCTTGTATTTCCAAAGTGGCGCTTACATATCGCAACCAAGGTTGGCGCTCT 5896
```

```
QY      781 GTTGAGACATCTGGTGTGTTGCTCAGGAGGAAGCCACTCCGGTCTTGTCTATGTTATGCATA 840
Db      5897 GTTGAGACCTCCGGGTGCGGAGGAGGAGGCCACTCTGTGTTCTGTCAATGCTCTGCATA 5956
QY      841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTTG 900
Db      5957 CATGGCTCACTGTAAATCTTATATACTAATACACCTATACCGGTGCGCTTGGCTGTGTTG 6016
QY      901 GACTTTGCTTATAGAGCTTGAGTTTGCATCTCACACCTGTAAACCAATACACCTGTG 960
Db      6017 GACTTTGCGCTCGAATCTTGAGTTCCGCAACCTCACCCCCGGTAATACCAATACCGGGTC 6076
QY      961 TCCCGTTACTCCAGCACTCTCTGTCACCTCC--GCCCGAGGGGCGGACGGGAGTCGGGAG 1017
Db      6077 TCGCGTTACTCCAGCACTGCCCTGTCACCGCTTCCGTCGCGGTGCAGATGGGAGTCGGGAG 6136
QY      1018 CTGACCAACCTGCGAGCCACAGGTTTCATGAAGATCTCACTTTTACCGGCTTATATGGG 1077
Db      6137 CTCACCAACCGCTGCTACTCGCTTCATGAAGGACCTCTATTTTACTAGTACTAATGTT 6196
QY      1078 GTAGGTGAAGTCGCGCGGATAGCTCTAAACATTACTTAACTTCTGCTGACACGCTCCTC 1137
Db      6197 GTTGTGAGATCGCGCGGAGTAGCGCTTACCCTGTTTAACTTGTCTGACACCTTGTCT 6256
QY      1138 GCGGGCTCCCGACAGAAATTAATTCGTCGCGTGGCGGCAACTGTTTATTTCCCGCCCG 1197
Db      6257 GCGGCTCTACCCACAGAAATTAATTCGTCGCGTGGTGGCAGCTGTTCTACTCTCGCCCG 6316
QY      1198 GTTGTCTAGCCAAATGCGGAGCCAACTGTAAGCTCTATACATCAGTGGAGATGCTCAG 1257
Db      6317 GTCGTCTCAGCAATGGCGAGCGGACTGTTAAGCTGTATACATCTGTGGAGATGCTCAG 6376
QY      1258 CAGGATAAGGGTGTGCTATCCCCACGATATCGATCTTGGTGAATTCGCGTGTGTCATT 1317
Db      6377 CAGGATAAGGGTATTGCAATCCGCAATGACATCGACTCGGGGAATCCCGTGTAGTATT 6436
QY      1318 CAGGATTATGACAAACAGCATGAGCAGATCGGCGCCACCCCGTGCCTGCGGCATCTCGG 1377
Db      6437 CAGGATTATGACAAACAAATGAGCAGGACCGACCCAGACACTTCCCGCAGCCCATCGCGT 6496
QY      1378 CTTTTTCTGTTCTCGAGCAATGATGTACTTTGCTGTCTCTGCTGCTGCTGCTGCTGCTG 1437
Db      6497 CTTTTTCTGCTCTCGAGCTAAACGATGTGCTTGGCTTCTCTCAACCGTGGCGAGTAT 6556
QY      1438 GACCACTCCACTTACGGGTCTGCAACTGGCCCGGTTTATATCTCGACAGACGCTGATTTG 1497
Db      6557 GACCACTCCACTTACGGCTCTTCGACCGGCCAGTCTATGTCTCTGACTCTGTGACCTTG 6616
QY      1498 GTGAATGTTGCGACTGGCGCGCAGCGCTGAGCCGATCGCTTGAATCGGTTCCAAAGTCACC 1557
Db      6617 GTTAATGTTGCGACCGCGCGCAGGCGGTTGCCGCTCACTCGACTGGACCAAGGTCACA 6676
QY      1558 CTCGACGGGCGGCCCCCTCCGACTGTTGAGCAATATTTCCAAAGCAATTTCTTGTGCTCCCG 1617
Db      6677 CTTGATGTTGCGCCCCCTTCCACCACTCCAGCAGTATTTCAAAGACCTTCTTTGCTTGGCG 6736
QY      1618 CTTGCTGGCAAGCTCTCTCTTTTGGGAGGCGGCAACAACAAAAGCAGGTTATCTCTATAAT 1677
Db      6737 CTCGCGGTAAAGCTCTCTCTTTTGGGAGGCGAGGAACCTATAAGCCGGTACCTTATAAT 6796
QY      1678 TATAATCTACTGCTAGTGACACGATTTCTGAATGGAATAATGCTGCGCGCCATCGGGTCGCC 1737
Db      6797 TATAAACCACTGCTAGTGACCAACTGCTCGTTGAGAAATGCCGCTGGGATCGGGTGTCT 6856
QY      1738 ATTTCAACCTATACACAGGCTTGGGCGCGTCCGCTCGCAATTTCTCGCGCGCGGTT 1797
Db      6857 ATTTCCACCTTACACTACTAGCTGGGTGCTGGCCCCGCTCTCTATTTCCGCGGTTGCTGTT 6916
QY      1798 TTGGCTCCACGCTCCGCCCTTGGCTCTGCTGGAGGATACATTTTGAATATCCGGGCGGGCG 1857
Db      6917 TTAGCCCCCACCTCTGTGCTAGCAATTTGCTTGAGGATACCATGACTACCTGCGCGCGCC 6976
QY      1858 CACACATTTGATGACTTCTGCCCTGAAATGCGCGCTTTAGGCTTCCAGGCTTGTGCTTTC 1917
```

Db	6977	CATACTTCGATGACTTCTGCCCGAGTGC	CGCCCTTGCCCTCCAGGGTGTGCTTTT	7036
Qy	1918	CAGTCAACTGTCGCTGAGCTCCAGCGCTTAAAGCTAAAGTGGGTAAAACTCCGGGAGTTG	1977	
Db	7037	CAGTCTACTGTCGCTGAGCTTCAGCGCTTAAAGTGAAGTGGGTAAAACTCCGGGAGTTA	7096	
Qy	1978	TAG 1980		
Db	7097	TAG 7099		
RESULT 14				
LOCUS	ARL139826	7168 bp	DNA	linear PAT 16-JUN-2001
DEFINITION	Sequence 4 from patent US 6207416.			
ACCESSION	ARL139826			
VERSION	ARL139826.1	GI:14482322		
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 7168)			
AUTHORS	Tearev, S.A., Emerson, S.U. and Purcell, R.H.			
TITLE	Recombinant proteins of a Pakistani strain of hepatitis E and their use in diagnostic methods and vaccines			
JOURNAL	Patent: US 6207416-A 4 27-MAR-2001;			
FEATURES	Location/Qualifiers			
source	1..7168			
	/organism="unknown"			
ORIGIN	/mol_type="unassigned DNA"			
Query Match 69.6%; Score 1378.2; DB 6; Length 7168;				
Best Local Similarity 81.3%; Pred. No. 0;				
Matches 1612; Conservative 0; Mismatches 368; Indels 3; Gaps 1;				
Qy	1	ATGGCCCTTAGGCTCTTTTGTCTGTTTTCCTCTTGTTCCTGCTATGTTGCCGCGCCA	60	
Db	5147	ATGGCCCTCGGCTATTTGCTGTTGCTCTCATGTTTCTGCTATGCTGCCGCGCCA	5206	
Qy	61	CCGACCGGTGACCGCTGCGCCCGCTGCTGGCGGCGGAGCGCGGTACCGGCGGTGCT	120	
Db	5207	CCGCGCGGTGACGCGCTGCGCCCGCTGCTGGCGGCGGAGCGCGGTTCGGCGGTGCT	5266	
Qy	121	TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC	180	
Db	5267	TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC	5326	
Qy	181	CCCTTTGCCCGACAGGTGCGGTGCGGTGCGGTGTGAGCTTCGCCCTTCGCCAACAGCC	240	
Db	5327	CCCTTGCCCGCCGATGTACCGCTGCGCGCGGGCTGGACCTCGTGTGCCAACCCGCC	5386	
Qy	241	CGCGCACTTGGCTCCACTTGGCGAGATCAGGCCCGCGCCCTTCGCTGCCTCCGTCGC	300	
Db	5387	CGACCACTCGGCTCCGCTTGGCGGTGACAGGCCCGCGCCCTTCGCTGCCTCCGTCGC	5446	
Qy	301	CGACCTGCCACAGCGGGGCTGCGGCGCTGACGGCTGTGGCGCTGCCCATGACCTCA	360	
Db	5447	AGACCTACACAGCTGGGCGCGCGCTAACCGCGGTGCTCCGCGCCATGACCCCG	5506	
Qy	361	CCGCTCCGCAAGTTGATTCGCGGTGCAATTTACGCCCGCAGTATAATTTGTCTACT	420	
Db	5507	CCAGTGGCTGATGTTGATCTCCGCGGCGCATCTCTGCGCGGCGAGTAACTATCAACA	5566	
Qy	421	TCACCCCTGACATCCTCTGTGGCTCTGCGACTAAATTTAGTCTCTGTATGAGCCCTCT	480	
Db	5567	TCCTCCCTCTACCTCTTCCGTTGGCCACCGGACAAATTTGGTTCTTTACGCGCTCTCT	5626	
Qy	481	AATCCGCTCTGCGGTGAGGAGCGGTACTTAATCACTACATATGCGCCACAGAGGCTCC	540	
Db	5627	AGCCCGCTTCTACCCCTCCAGGAGGCAACCACTACTCATATAATGGCTACAGAGCTTCT	5686	

Db	6767	CTCCGGGTAAAGCTCTCTCTTTTGGGAGGAGGAACTACTAAAGCCGGTACCTTTATAAT	6826	QY	301	CGACCTGCGACAGCCGGGCTGCGCGCTGAGCGCTGTGCGCCTGCCCATGACACTCA	360
QY	1678	TATAATACCTAGCTAGTACAGAGTCTGATTTGAAATAGTGTGCGGCGCATCGGTGCGC	1737	Db	5447	AGACCTTACCACAGCTGGGSCCGCGCTAAACGGGGTGCCTCCGGGCCCATGACACCCCG	5506
Db	6827	TATAACCACTAGCTAGTGACCACTGCTCGTTGAGAAATGCGCTGGGCATCGGGTTGCT	6886	QY	361	CCCGTCCCGGACGTTGATTTCTCGCGTGCAATTTCTACGCGCGCAGTATATTTGTCTACT	420
QY	1738	ATTTCNAACCTATACACAGGCTTTGGGCGCGTCCGGTCGCCATTTCTGGGCGCGGTT	1797	Db	5507	CCAGTGCCTGATGTTGACTCCCGCGCGCCATCTCTGCGCGCAGTATAACCTATCAACA	5566
Db	6887	ATTTCACCTACACTAGCTAGCTGGGTGCTGCGCCGCTCTCTATTTCCGCGGTGCTGTT	6946	QY	421	TCAACCCCTGACATCCCTCTGTGGCCTCTGGCACAATAATTTAGTCTCTGTATCGAGCCGCCCTT	480
QY	1798	TTGGGTCCACAGCTCCGCCCTGGCTCTGCTGGAGGATACATTTTGAATATCCGGGCGGGCG	1857	Db	5567	TTCTCCCTCACCTCTTTCCGTGGCCACCGGCACAAATTTGGTCTTTTACGCGCTCTCTTT	5626
Db	6947	TTAGCCCCCACTCTGTGTAGCAATGCTTGAGGATACCATGGACTACCTGCGCGCC	7006	QY	481	AATTCGCGCTCTCTCGCTGCGAGACGCTACTATACTACATATATGGCCACAGAGGCTCC	540
QY	1858	CACACATTTGATGACTTTGCGCTGAAATGCGCGGCTTTAGGCTTCAGGGTTGTGTTTC	1917	Db	5627	AGCCCGCTTCTACCCCTCCAGGACGSCACCACTACTATATAATATGGCTACAGAAGCTTCT	5686
Db	7007	CATACTTTGATGACTTTCTGCGGAGTGCAGCCCTTTGGCTCCAGGGTTGTGCTTTT	7066	QY	541	AATTAATGACAGTACCGGTTGCGCGCTACTATCCGTTACCGGCTTACCGGCTTAGTGCCTAT	600
QY	1918	CAGTCAACTGTCGCTGAGCTCCAGCGCTTAAAGTTTAAAGTGGTAAAACTCGGGAGTTG	1977	Db	5687	AATTAATGCCAGTACCGGTTGCTGTGCCCAATTTGCTACCGCCGCTGGTCCCAAC	5746
Db	7067	CAGTCTACTGTCCTGAGCTTACGCGCTTTAGATGAAGTGGTAAAACTCGGAGTTA	7126	QY	601	GCAGTTGAGGCTATGCTATATCCATTTCTTTCTGGCCTCAAAACACCAACCCCTACA	660
QY	1978	TAG 1980		Db	5747	GCTGTTGGTGGCTAGCTATCTCCATTTGCTTCTGGCCACAGCACACACCCCGAG	5806
Db	7127	TAG 7129		QY	661	TCTGTTGACATGAATTCANTTCTTCCACTGATGTGAGGATTTCTGTTCAACTGGCAT	720
RESULT 15				Db	5807	TCCGTTGACATGAATTCATAAATTAACCTCGACCGATGTCCGTATTTTACTCCAGCCG	5866
AR167470				QY	721	GCATCTGAATTTGCTATCCCAAGCAGCGCTTCACTACCGCAATCAAGTTGGCGCTCG	780
LOCUS	AR167470	7168 bp	DNA	linear	5867	GCTCCGAGCTTTGTTATTCCTCAAGTGAAGCGCTTACACTATCGCAACCAAGTTGGCGCTCT	5926
DEFINITION	Sequence 4 from patent US 6287759.			QY	781	GTTGAGACATCTGCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	840
ACCESSION	AR167470			Db	5927	GTTGAGACCTCCGGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	5986
VERSION	AR167470.1	GI:17903252		QY	841	CATGGCTCTCGAGTAACTCTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTACTG	900
KEYWORDS	Unknown.			Db	5987	CATGGCTCACTGTAATTTCTTATATACTATAACCTATACCCCTATACCGGTGCCCTTGG	6046
SOURCE	Unknown.			QY	901	GACTTTGCTTAGAGCTTGTGATTTTCCATCTTCCACCTGTAAACCAATACACCTGTG	960
ORGANISM	Unclassified.			Db	6047	GACTTTGCGCTCGAATTTGAGTTTCCGCAACCTCACCCCGGTAAATACCAATACCGG	6106
REFERENCE	1 (bases 1 to 7168)			QY	961	TCCCGTTACTCCAGCACTGCTGCTCACTCC--GC0CCGAGGGGCGCAGCGGAGTGGGAG	1017
AUTHORS	Tsarev,S.A., Emerson,S.U. and Purcell,R.H.			Db	6107	TCCGCTTACTCCAGCACTGCCGCTCACCGCTTCGTCGCGGTGAGATGGGAGTGGGAG	6166
TITLE	Recombinant proteins of a Pakistani strain of hepatitis E and their			QY	1018	CTGACCAACTGACGCCACCGGTTTCATGAAGAATCTCCACTTTTACCGGCTTAATGGG	1077
USE IN DIAGNOSTIC METHODS AND VACCINES	use in diagnostic methods and vaccines			Db	6167	CTCACCACCAAGGCTGCTACTCGCTTCATGAAGGAGCTCTATTTTACTAGTACTAATGGT	6226
JOURNAL	Patent: US 6287759-A 4 11-SEP-2001;			QY	1078	GTAGGTGAAGTCGGCGCGGATAGCTTAACATTAACCTTAACTCTGTCACAGCTCCTC	1137
FEATURES	Location/Qualifiers			Db	6227	GTTGGTGAGATCGCGCGGATAGCGCTTACCCCTGTTTAACTTGTGACACCTTGGCT	6286
source	1..7168			QY	1138	GGCGGCTCCCGACAGAAATTAATTTCTGCGGTGGGCGGCAACTGTTTTATTTCCGCGCG	1197
ORIGIN	/organism="unknown"			Db	6287	GGCGGCTTACCGACAGAAATTAATTTCTGCGGTGGTGGCCAGCTGTTTCTACTCTCG	6346
Query Match	69.6%; Score 1378.2; DB 6; Length 7168;			QY	1198	GTTGCTCAGCNAATGGGAGCCAACTGTAAGCTCTATACATCAGTGGAGAAATGCTCAG	1257
Best Local Similarity	81.3%; Pred. No. 0;			Db	6347	GTGCTCTCAGCCAAATGGCGCGGCTGTTAAGCTGTATACATCTGTGGAGAAATGCTCAG	6406
Matches 1612; Conservative	0; Mismatches 368; Indels 3; Gaps 1;			QY	1258	CAGGATAAGGGTGTGCTATCCCCACGATATCGATCTTTGGTGAATTCGGGTGGTTCATT	1317
QY	1	ATGCGCCCTAGGCTCTTTTGGTGTGTTCTTCTGCTTGTGCTATGTTGCGCGGCA	60	Db	6407	CAGGATAAGGGTGAATTCGAATTCGCGATGATCGACTCGGGGAATTCGCGTGTAGTATT	6466
Db	5147	ATGCGCCCTCGGCTATTTGCTGTTGCTCTCATGTTTCTGCTATGCTGCGCGGCA	5206	QY	1318	CAGGATTTATGACAAACAGCATGAGCAGGATTCGCGCCCAACCCCGCTCGCGCTCGCATCTCG	1377
QY	61	CCGACCGGTACGCGTCTGCGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG	120	Db	6467	CAGGATTTATGACAAACCAACATGAGCAGGAGCCGACCGACACCTTTCCCGACCCCATCGCT	6526
Db	5207	CGCGCCGGTACGCGCTGTCGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG	5266	QY	1378	CCTTTTCTGTTCTCCGAGCAAAATGATGTACTTTGGTGTGCTCCCTCACTGAGCGGAGTAT	1437
QY	121	TTCTGGGGTGACCGGCTTGATTTCTAGCCCTTCGCAATCCCTTATATTCATCCAAAC	180				
Db	5267	TTCTGGGGTGACCGGCTTGATTTCTAGCCCTTCGCAATCCCTTATATTCATCCAAAC	5326				
QY	181	CCCTTTGCCCGACAGTTCGCGCTGCGGTCTGGAGCTCGCCTTCGCCCAACAGCC	240				
Db	5327	CCCTTTGCCCGCGATGTACCGCTGCGGCGGCGGCTGGACCTGCTGTTCCCAACCGCC	5386				
QY	241	CGGCGCTTGGCTCCACTTGGCGAGATCAGGCCCGCCCTCCGCTGCTCCGCTGCG	300				
Db	5387	CGACCACTCGGCTCCGCTTGGCGTGACCGGCCAGCGCCCGCGCTGCTCAGCTCGT	5446				

Db	6527	CCCTTTTCTGTCTCCGAGCTAACGATGTGCTTTGGCTTTCTCTCACCGTCCGAGTAT	6586
Qy	1438	GACCAGTCCACTTACGGGTGCTCAACTGGCCCGGTTTATATCTCGGACAGCGTGACTTTG	1497
Db	6587	GACCAGTCCACTTACGGCTCTTCGACGGCCAGTCTATGCTCTGACTCTGTGACCTTG	6646
Qy	1498	GTGAATGTTGGGACTGGGGCGCAGCCGTAGCCCGATGCGTTGACTGGTCCAAAGTCACC	1557
Db	6647	GTTTAATGTTGGCAGCGGCGCAGCCGTTGCCCGGTCACTCGACTGGACCAAGGTCACA	6706
Qy	1558	CTCGACGGGGGCCCTCCGACCTGTTGAGCAATATTCAAAGACATTCTTTGTGCTCCCC	1617
Db	6707	CTTGATGGTCGCCCCCTTCCACCATCCAGCAGTATTCAAAGACCTTCTTTGCTGCCG	6766
Qy	1618	CTTCGTGGGAAGCTCTCGTTTTGGGAGCCCGGCAACAACAAAGCAGGTTATCCTTATAAT	1677
Db	6767	CTCCGCGGTAAAGCTCTCCTTTGGGAGGCGAGGAATACTAAGCCGGGTACCTCTATAAT	6826
Qy	1678	TATAAATACTACTGCTAGTGACCAGATTCTGATTGAAAATGTCGCCGCCATCGGGTCGCC	1737
Db	6827	TATAACACCACCTGCTAGTGACCAACTGCTCGTTGAGAATCGCGCTGGCATCGGGTGTCT	6886
Qy	1738	ATTTCACCTATTACCACAGGCTTGGGGCCGGTCGGCTCGGCATTTCTCGGCCCGCGGTT	1797
Db	6887	ATTTCACCTACACTACTAGCTGGGTGCTGGCCCGCTCTATTTCCGGGGTGTGCTGTT	6946
Qy	1798	TTGGCTCCAGCTCCGGCCCTGGCTCTGCTGGAGGATACTTTTGATTATCGGGCGGGCG	1857
Db	6947	TTAGCCCCACACTCTGTGCTAGCATTTGTTGAGGATACCATGGACTACCTGCGCGGCC	7006
Qy	1858	CACACATTTGATGACTTCTGCGCTTGAATGCGCGCTTTAGCCCTCCAGGGTGTGCTTTC	1917
Db	7007	CATATCTTCATGACTTCTGCGCGGAGTGCGCCGCCCTTGCCCTCCAGGGTGTGCTTTT	7066
Qy	1918	CAGTCAACTGTCGCTGAGCTCCAGCGCTTAAAGCTTAAGGTGGGTAAACCTCGGAGCTG	1977
Db	7067	CAGTCTACTGTCGCTGAGCTTCAGCGCTTAAAGTGAAGGTGGGTAAACCTCGGAGTTA	7126
Qy	1978	TAG 1980	
Db	7127	TAG 7129	

Search completed: August 21, 2004, 15:15:06
Job time : 7637.72 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 08:40:47 ; Search time 764.421 Seconds
(without alignments)
11003.666 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096-
Perfect score: 1980
Sequence: 1 ATGGCGCCTAGGCCTCTTTT.....GTAAACTCGGAGTTGTAG 1980

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002s.*
- 7: geneseqn2003as.*
- 8: geneseqn2003bs.*
- 9: geneseqn2003cs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1980	100.0	1980	2	AAT96960	Aat96960 Hepatitis
2	1980	100.0	2100	2	AAQ47130	Aaq47130 HEV third
3	1980	100.0	7171	2	AAQ14413	Aaq14413 Composite
4	1980	100.0	7171	2	AAV66322	Aav66322 ET-NANB (
5	1980	100.0	7171	2	AAV54730	Aav54730 Composite
6	1980	100.0	7171	3	AAQ99260	Aaq99260 Hepatitis
7	1980	100.0	7171	4	AAF83496	Aaf83496 Mexican s
8	1980	100.0	7171	6	AAU50387	Aau50387 Hepatitis
9	1980	100.0	7171	9	ADD24378	Add24378 Hepatitis
10	1980	100.0	7180	9	ADE06698	Ade06698 Hepatitis
11	1978.4	99.9	2100	2	AAQ46814	Aaq46814 Mexico st
12	1957.4	98.9	2058	2	AAT27108	Aat27108 Hepatitis
13	1621.4	81.9	1647	2	AAT27110	Aat27110 Hepatitis
14	1378.2	69.6	1983	2	AAV61688	Aav61688 Hepatitis
15	1378.2	69.6	7158	2	AAT27394	Aat27394 Hepatitis
16	1378.2	69.6	7168	2	AAQ45197	Aaq45197 HEV strai
17	1378.2	69.6	7168	2	AAV71604	Aav71604 Hepatitis
18	1378.2	69.6	7194	2	AAV61690	Aav61690 Hepatitis
19	1378.2	69.6	7204	8	ADA50064	Ada50064 Hepatitis
20	1378.2	69.6	7204	8	ADA50063	Ada50063 SK-HEV-2
21	1376.6	69.5	7204	8	ADA50062	Ada50062 SK-HEV-3
22	1376.6	69.5	7204	8	ADA50065	Ada50065 Hepatitis
23	1371.8	69.3	1984	2	AAQ86592	Aaq86592 Hepatitis

24	1371.8	69.3	1990	6	ABL60053	Abi60053 Hepatitis
25	1371.8	69.3	1990	9	ADB97808	Adb97808 HEV ORF2
26	1371.8	69.3	2054	4	AAQ01448	Aas01448 Hepatitis
27	1368.6	69.1	1983	2	AAT96959	Aat96959 Hepatitis
28	1368.6	69.1	2049	2	AAQ27107	Aat27107 Hepatitis
29	1368.6	69.1	2094	2	AAQ47129	Aaq47129 HEV third
30	1368.6	69.1	2094	2	AAQ46813	Aaq46813 Burma str
31	1368.6	69.1	7195	2	AAV66321	Aav66321 ET-NANB (
32	1368.6	69.1	7195	2	AAV54729	Aav54729 DNA seque
33	1368.6	69.1	7195	3	AAV9259	Aaa9259 Hepatitis
34	1368.6	69.1	7195	4	AAF83495	Aaf83495 ET-NANB v
35	1368.6	69.1	7195	6	AAU50386	Aau50386 Hepatitis
36	1363.8	68.9	7195	9	ADD24374	Add24374 Hepatitis
37	1361.4	68.8	1989	6	ABL60054	Abi60054 Hepatitis
38	1361.4	68.8	1989	9	ADB97809	Adb97809 HEV relat
39	1355.8	68.5	7195	2	AAQ14412	Aaq14412 Forward s
40	1319.6	66.6	7256	9	ADE06693	Ade06693 Hepatitis
41	1315.2	66.4	7277	2	AAZ00267	Aaz00267 HEV-US2 O
42	1315.2	66.4	7277	2	AAZ00266	Aaz00266 HEV-US2 f
43	1314	66.4	7228	9	ADE06696	Ade06696 Hepatitis
44	1313.2	66.3	7233	9	ADE06694	Ade06694 Hepatitis
45	1301.2	65.7	7229	9	ADE06695	Ade06695 Hepatitis

ALIGNMENTS

RESULT 1
AAT96960
ID AAT96960 standard; DNA; 1980 BP.
XX
AC AAT96960;
XX
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 26-FEB-1998 (first entry)
XX
DE Hepatitis E virus Mexico strain ORF2.
XX
KW Hepatitis E virus; Burma; Mexico; immunoassay; peptide antigen; antibody;
KW diagnosis; HEV; ds.
XX
OS Hepatitis E virus; - Mexico strain.
XX
FH Key Location/Qualifiers
FT CDS 1..1980
FT /*tag= a
FT /product= "ORF2_protein"

XX US5686239-A.
XX
XX 11-NOV-1997.
XX
XX 09-MAY-1994; 9AUS-00240049.
XX
XX 17-JUN-1988; 88US-00208997.
XX 11-APR-1989; 85US-00336672.
XX 16-JUN-1989; 89US-00367486.
XX 13-OCT-1989; 89US-00420921.
XX 05-APR-1990; 90US-00505888.
XX 05-APR-1991; 91US-00681078.
XX 17-JAN-1992; 92US-00822335.
XX 20-APR-1992; 92US-00870985.
XX 01-MAY-1992; 92US-00876941.
XX
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX
XX Tam AW, Reyes GR, Yarbough PO;
XX
XX WPI; 1997-558132/51.
XX P-PSDB; AAW35827.
XX
XX Diagnosis of hepatitis E virus Burma and Mexico strain infection - by

PT immunoassay with hepatitis E virus peptide antigens.

PS Disclosure; Col 25-28; 36pp; English.

XX A method has been developed for detecting hepatitis E virus (HEV) antibodies (Ab). The method comprises: (a) reacting a serum sample with a HEV peptide antigen; and (b) examining the peptide for the presence of bound Ab, where the presence of bound Ab indicates the presence of HEV Ab. The present sequence encodes the protein from the open reading frame, ORF2, from HEV Mexico strain. The method can be used to diagnose infection with the enterically transmitted non-A/non-B viral hepatitis agent HEV, specifically the HEV Burma and Mexico strains. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 1980 BP; 364 A; 624 C; 493 G; 499 T; 0 U; 0 Other;

Query Match 100.0%; Score 1980; DB 2; Length 1980;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTAGGCTCTTTGCTGTTGTTTCCTCTGTTCTGCTATGTTGCCGGCCA 60

DB 1 ATGGCCCTAGGCTCTTTGCTGTTGTTTCCTCTGTTCTGCTATGTTGCCGGCCA 60

QY 61 CCGACCGGTGAGCGGTCTGGCCCGCTGCTGGGGCGCGAGCGGTACCGCGGTGGT 120

DB 61 CCGACCGGTGAGCGGTCTGGCCCGCTGCTGGGGCGCGAGCGGTACCGCGGTGGT 120

QY 121 TTCTGGGGTGACGGGTGATTTCTAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180

DB 121 TTCTGGGGTGACGGGTGATTTCTAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180

QY 181 CCCTTTGCCCGACAGCTTCGGCTGCTGGGTCTGACCTCGCTTCGCCAACCGCC 240

DB 181 CCCTTTGCCCGACAGCTTCGGCTGCTGGGTCTGACCTCGCTTCGCCAACCGCC 240

QY 241 CGGCCACTTGGCTTCCACTTGGCGAGATCAGGCCCGCCCTCGCTCGCTCCCGTCGC 300

DB 241 CGGCCACTTGGCTTCCACTTGGCGAGATCAGGCCCGCCCTCGCTCGCTCCCGTCGC 300

QY 301 CGACCTGCCACAGCGGGGCTGCGGCTGACGGCTGTGGGCTGTGCCCATGACACTCA 360

DB 301 CGACCTGCCACAGCGGGGCTGCGGCTGACGGCTGTGGGCTGTGCCCATGACACTCA 360

QY 361 CCGTCCCGGAGCTTGAATTTCTCGCGTGCAATTTCTACGCGCGCAGTAAATTTGCTACT 420

DB 361 CCGTCCCGGAGCTTGAATTTCTCGCGTGCAATTTCTACGCGCGCAGTAAATTTGCTACT 420

QY 421 TCACCCCTGACATTCCTCTGTGGCTCTGGCACTAATTTAGTCTGTATGACGCCCTCTT 480

DB 421 TCACCCCTGACATTCCTCTGTGGCTCTGGCACTAATTTAGTCTGTATGACGCCCTCTT 480

QY 481 AATCGCCTCTGCGCTGAGAGCGTACTAATCTACATATGCGCCACAGAGGCTCC 540

DB 481 AATCGCCTCTGCGCTGAGAGCGTACTAATCTACATATGCGCCACAGAGGCTCC 540

QY 541 AATTATGACAGTACCGGTTGCGCGCTACTATCCGTTACGGCCCTAGTGCCTAAT 600

DB 541 AATTATGACAGTACCGGTTGCGCGCTACTATCCGTTACGGCCCTAGTGCCTAAT 600

QY 601 CGAGTTGGAGGCTATGCTATATCCATTTCTTCTGGCTCAAAACCAACCCCTACA 660

DB 601 CGAGTTGGAGGCTATGCTATATCCATTTCTTCTGGCTCAAAACCAACCCCTACA 660

QY 661 TCTGTTGACATGAATTCATTCCTACTGATGTCAGGATTTCTGTTCAACCTGSCATA 720

DB 661 TCTGTTGACATGAATTCATTCCTACTGATGTCAGGATTTCTGTTCAACCTGSCATA 720

QY 721 GCATCTGAATTTGGTCAATCCCAAGCAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 780

DB 721 GCATCTGAATTTGGTCAATCCCAAGCAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 780

QY 781 GTTGAGACATCTGGTGTCTGTAGGAGGAAGCACCTCGGTCTTGTCTGTATGATGATA 840

DB 781 GTTGAGACATCTGGTGTCTGTAGGAGGAAGCACCTCGGTCTTGTCTGTATGATGATA 840

QY 841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCTTGGCTACTG 900

DB 841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCTTGGCTACTG 900

QY 901 GACTTTGCCCTTAGAGCTTGAGTTTGCATCTCACCACTCTAACAACCAATACACGTG 960

DB 901 GACTTTGCCCTTAGAGCTTGAGTTTGCATCTCACCACTCTAACAACCAATACACGTG 960

QY 961 TCCCGTTACTCCAGCACTCTCGTCACTCCGCCGAGGGCCGACGGGAGCTGCGGAGCTG 1020

DB 961 TCCCGTTACTCCAGCACTCTCGTCACTCCGCCGAGGGCCGACGGGAGCTGCGGAGCTG 1020

QY 1021 ACCAACAAGTGCAGCCACAGGTTTCATGAAAGATCTCCACTTTACCGGCCCTTAAATGGGGTA 1080

DB 1021 ACCAACAAGTGCAGCCACAGGTTTCATGAAAGATCTCCACTTTACCGGCCCTTAAATGGGGTA 1080

QY 1081 GGTGAAGTCCGCCGCGGGATAGCTCTAACATTTAACTTTGCTGACACGCTCTCTCGGC 1140

DB 1081 GGTGAAGTCCGCCGCGGGATAGCTCTAACATTTAACTTTGCTGACACGCTCTCTCGGC 1140

QY 1141 GGGCTCCCGACAGAAATTAATTTCTGGCTGGCGGCAACTGTTTTATTTCCGCGCCGGTT 1200

DB 1141 GGGCTCCCGACAGAAATTAATTTCTGGCTGGCGGCAACTGTTTTATTTCCGCGCCGGTT 1200

QY 1201 GTCTAGCCAAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGATGCTCAGCAG 1260

DB 1201 GTCTAGCCAAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGATGCTCAGCAG 1260

QY 1261 GATAAGGTTGTTGCTATCCCCACCATATCGATCTTGGTGATTTGGGTGCTGCTCATTCAG 1320

DB 1261 GATAAGGTTGTTGCTATCCCCACCATATCGATCTTGGTGATTTGGGTGCTGCTCATTCAG 1320

QY 1321 GATTATGACCAACAGCAGATGCGGCCACCCCGCTGCGCCATCTCGGCCCT 1380

DB 1321 GATTATGACCAACAGCAGATGCGGCCACCCCGCTGCGCCATCTCGGCCCT 1380

QY 1381 TTTTCTGTTCTCGAGCAAAATGATGTAATTTGGTGTCTCTCCTCCTCAGCAGCGATGAC 1440

DB 1381 TTTTCTGTTCTCGAGCAAAATGATGTAATTTGGTGTCTCTCCTCCTCAGCAGCGATGAC 1440

QY 1441 CAGTCCACTTACGGTCTGCTCAACTGGCCGGTTTATATCTCGACAGCGCTGACTTTGGTG 1500

DB 1441 CAGTCCACTTACGGTCTGCTCAACTGGCCGGTTTATATCTCGACAGCGCTGACTTTGGTG 1500

QY 1501 AATGTTGCACTGGCGCGAGCCCGTAGCCCGATCGCTTGAAGTCCAAAGTCAACCTC 1560

DB 1501 AATGTTGCACTGGCGCGAGCCCGTAGCCCGATCGCTTGAAGTCCAAAGTCAACCTC 1560

QY 1561 GACGGCGCCCTCTCCGACTGTTGAGCAATATCCAGACATTTCTTGTCTCCCGCTT 1620

DB 1561 GACGGCGCCCTCTCCGACTGTTGAGCAATATCCAGACATTTCTTGTCTCCCGCTT 1620

QY 1621 CGTGGCAAGCTCTCTTTTGGAGCGCGGCAAAAGAGAGGTTATCTTATAATTAAT 1680

DB 1621 CGTGGCAAGCTCTCTTTTGGAGCGCGGCAAAAGAGAGGTTATCTTATAATTAAT 1680

QY 1681 AATACTACTGCTAGTGCAGATTTGATTTGAAATGTCGCGGCCATCGGGTCGCAAT 1740

DB 1681 AATACTACTGCTAGTGCAGATTTGATTTGAAATGTCGCGGCCATCGGGTCGCAAT 1740

QY 1741 TCAACTATACCAACAGGCTTGGCGCGGTCCGCTCGCATTTCTGCGCGCGGGTTTG 1800

DB 1741 TCAACTATACCAACAGGCTTGGCGCGGTCCGCTCGCATTTCTGCGCGCGGGTTTG 1800

QY 1801 GCTCCACGCTCCGCCCTGCTCTGCTGGAGGATCTTTTGAATTCGCGGGCGGGCGCAC 1860

DB 1801 GCTCCACGCTCCGCCCTGCTCTGCTGGAGGATCTTTTGAATTCGCGGGCGGGCGCAC 1860

QY 1861 ACATTTGATGACTTCTGCGCCTGAATGCGCGCTTTAGGCTCCAGGGTTGTGCTTCCAG 1920

|||||

Db 1861 ACAATTGATGACTTCTGCCCTGAATGCGCGCTTTAGGCTCCAGGGTTGCTTTCCAG 1920

QY 1921 TCAACTGTCGTGAGCTCCAGGCTTAAAGTTAAGTGGGTAAACTCGGAGTTGTAG 1980

Db 1921 TCAACTGTCGTGAGCTCCAGGCTTAAAGTTAAGTGGGTAAACTCGGAGTTGTAG 1980

RESULT 2

AAQ47130

ID AAQ47130 standard; DNA; 2100 BP.

XX AC AAQ47130;

XX DT 24-OCT-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 11-JAN-1994 (first entry)

XX DE HEV third and second ORFs.

XX KW Enterically transmitted non-A non-B hepatitis; ET-NANB; vaccine; ss.

XX OS Hepatitis E virus; Mexico strain.

XX FH Key Location/Qualifiers

FT CDS 5..376

FT FT /*tag= a

FT FT /label= ORF3

FT CDS 46..2100

FT FT /*tag= b

FT FT /label= ORF2

FT FT 275..376

FT FT /*tag= c

FT FT /label= 406.4-2

FT FT 718..2026

FT FT /*tag= d

FT FT /label= C2

FT FT 1045..2026

FT FT /*tag= e

FT FT /label= SG3

FT FT 1882..2026

FT FT /*tag= f

FT FT /label= 406.3-2

XX WO9314116-A1.

XX 22-JUL-1993.

XX 15-JAN-1993; 93WO-US000459.

XX 17-JAN-1992; 92US-00822335.

XX 01-MAY-1992; 92US-00876941.

XX (GENE-) GENELABS TECHNOLOGIES INC.

XX (USSH) US SEC DEPT HEALTH.

XX Reyes GR, Bradley DW, Tam AW, Carl M;

XX WPI; 1993-243144/30.

XX P-PSDB; AAR38786, AAR38787.

XX New immunogenic hepatitis E virus (HEV) peptide(s) - are from the ORF1, ORF2 and ORF3 regions of HEV, useful as a vaccine against HEV infection.

XX Disclosure; Fig 7; 48pp; English.

XX Immunogenic hepatitis E virus (HEV) peptides are selected from the ORF1, ORF2 and ORF3 regions of HEV. The peptides can be used in vaccines to prevent infection by HEV. The antibodies can neutralise and block HEV infection and can be used to prevent or treat HEV infection. The peptides and antibodies can also be used as diagnostic reagents. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS field)

XX SQ Sequence 2100 BP; 381 A; 659 C; 516 G; 544 T; 0 U; 0 Other;

Query Match 100.0%; Score 1980; DB 2; Length 2100;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTAGGCTCTTTGCTGTGTGTTCTCTCTGTTTCTGCTATGTTGCGCGGCA 60

Db ATGGCCCTAGGCTCTTTGCTGTGTGTTCTCTCTGTTTCTGCTATGTTGCGCGGCA 105

QY 61 CCGACCGGTTCAGCCGCTCTGGCGCGCGCTGTGGCGCGCGAGCGCGGTACCGCGCGGTGT 120

Db CCGACCGGTTCAGCCGCTCTGGCGCGCGCTGTGGCGCGCGAGCGCGGTACCGCGCGGTGT 165

QY 121 TTCTGGGTGACCGGGTTGATTCTCAGCCCTTGGCAATCCCTATATATTCATCAACCAAC 180

Db TTCTGGGTGACCGGGTTGATTCTCAGCCCTTGGCAATCCCTATATATTCATCAACCAAC 225

QY 181 CCCTTTGCCCGAGAGTTGCGCTGCGCTCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 240

Db CCCTTTGCCCGAGAGTTGCGCTGCGCTCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 285

QY 241 CGGCACCTTGGCTCCACTTGGCGAGATCAGGCGCCAGCGCCCTCCGCTGCGCTGCGCTGCG 300

Db CGGCACCTTGGCTCCACTTGGCGAGATCAGGCGCCAGCGCCCTCCGCTGCGCTGCGCTGCG 345

QY 301 CGACTGCCACAGCGCGGCTGCGCGCTGAGCGGTGCGCGCTGCGCGCTGCGCGCTGCGCG 360

Db CGACTGCCACAGCGCGGCTGCGCGCTGAGCGGTGCGCGCTGCGCGCTGCGCGCTGCGCG 405

QY 361 CCGCTCCCGGACGTTGATTCTCGGGTGCAATCTACCGCGCGCGCTGCAATTTGTTGTTACT 420

Db CCGCTCCCGGACGTTGATTCTCGGGTGCAATCTACCGCGCGCGCTGCAATTTGTTGTTACT 465

QY 421 TCACCCCTGACATCCTCTGTGGCCCTCTGGCACTAATTTAGTCTCTGTATGCGACCCCTTT 480

Db TCACCCCTGACATCCTCTGTGGCCCTCTGGCACTAATTTAGTCTCTGTATGCGACCCCTTT 525

QY 481 AATCGGCTCTGCGCTGCGGACGAGTACTAATACTACATTATGGCCACAGAGCCCTCC 540

Db AATCGGCTCTGCGCTGCGGACGAGTACTAATACTACATTATGGCCACAGAGSCCTCC 585

QY 541 AATTATGCACAGTACCGGTTGCGCGCTACTACTCCGTTACCGGCCCTTAGTGCTAAT 600

Db AATTATGCACAGTACCGGTTGCGCGCTACTACTCCGTTACCGGCCCTTAGTGCTAAT 645

QY 601 GCAGTTGGAGGCTATGCTATATCAATTTCTTCTGGCCCTCAAAACAACCAACCCCTACA 660

Db GCAGTTGGAGGCTATGCTATATCAATTTCTTCTGGCCCTCAAAACAACCAACCCCTACA 705

QY 661 TCTGTTGACATGAATTCATTACTTCCACTGATGTGAGGATTCCTGTTCAACCTGGCATA 720

Db TCTGTTGACATGAATTCATTACTTCCACTGATGTGAGGATTCCTGTTCAACCTGGCATA 765

QY 721 GCATCTGAATTGTTTCATCCAGCGGCTTCACTACCGCAATCAAGTTGGGCTCG 780

Db GCATCTGAATTGTTTCATCCAGCGGCTTCACTACCGCAATCAAGTTGGGCTCG 825

QY 781 GTTCAGACATCTGCTGTGTGAGGAGGAGCCACTCCGCTCTTGTCTATGATGATA 840

Db GTTCAGACATCTGCTGTGTGAGGAGGAGCCACTCCGCTCTTGTCTATGATGATA 885

QY 841 CATGGCTCTCCAGTAACTCCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 900

Db CATGGCTCTCCAGTAACTCCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 945

QY 901 GACTTTGCCCTTAGAGCTTGAGTTTCGCAATCTACCACTGTATACCAATACAGGTGTG 960

Db GACTTTGCCCTTAGAGCTTGAGTTTCGCAATCTACCACTGTATACCAATACAGGTGTG 1005

QY 961 TCCGTTTACTCCAGCACTGCTCGTCACTCCGCGGAGGCGCGGAGGACTCGGAGCTG 1020

XX WPI; 1998-582599/49.
XX Hepatitis E virus proteins - useful for diagnosis or vaccine production
XX the virus.
XX
XX Claim 4; Col 69-76; 47pp; English.
XX
XX The present sequence represents the composite Mexico strain sequence of
XX enterically transmitted non A non B hepatitis virus (ET-NANB) (hepatitis
XX E virus (HEV)). The specification describes an isolated protein which is
XX specifically immunoreactive with antibodies present in individuals
XX infected with HEV and encoded by a sequence contained in an open reading
XX frame (ORF) of an HEV genome. The genome has a sequence that is more than
XX 70% identical to the ORF1 sequence from Burma HEV isolate (AAV6321). The
XX protein is used as a vaccine and a diagnostic probe for ET-NANB. (Updated
XX on 25-MAR-2003 to correct PR field.) (Updated on 27-AUG-2003 to correct
XX OS field.)
XX
SQ Sequence 7171 BP; 1285 A; 2147 C; 1907 G; 1832 T; 0 U; 0 Other;
Query Match 100.0%; Score 1980; DB 2; Length 7171;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTAGGCTCTTTGGCTGTGTTCTCTTCTGCTATGTTGCCCGGCCA 60
DB 5117 ATGCGCCCTAGGCTCTTTGGCTGTGTTCTCTTCTGCTATGTTGCCCGGCCA 5176
QY 61 CCGACCGGTCAGCGGTCCTGGCCCGCTCGTGGCGCGCAGCGCGGTACCGCGGTGGT 120
DB 5177 CCGACCGGTCAGCGGTCCTGGCCCGCTCGTGGCGCGCAGCGCGGTACCGCGGTGGT 5236
QY 121 TTCTGGGGTGACCGGGTGATTTCTAGCCCTTCGGAATCCGCTATATTCATCAACCAAC 180
DB 5237 TTCTGGGGTGACCGGGTGATTTCTAGCCCTTCGGAATCCGCTATATTCATCAACCAAC 5296
QY 181 CCCTTTGCCCCAGACGTCGCGTGGCTGGGCTCGGACCTCGCCCTCGCCCAACAGCC 240
DB 5297 CCCTTTGCCCCAGACGTCGCGTGGCTGGGCTCGGACCTCGCCCTCGCCCAACAGCC 5356
QY 241 CGGCCACTTGGCTTCCACTTGGCGAGATCAGGCCAGCGCCCTCGCTCGCTCCCGTCG 300
DB 5357 CGGCCACTTGGCTTCCACTTGGCGAGATCAGGCCAGCGCCCTCGCTCGCTCCCGTCG 5416
QY 301 CGACCTGCCACAGCCGGGCTCGCGGCTGACGGCTGTGGCGCTCGCCATGACACTCA 360
DB 5417 CGACCTGCCACAGCCGGGCTCGCGGCTGACGGCTGTGGCGCTCGCCATGACACTCA 5476
QY 361 CCGTCCCGGACGTTGATTTCTCGCGTGCAATTTCTACGCGCCAGTATATTTGTTACT 420
DB 5477 CCGTCCCGGACGTTGATTTCTCGCGTGCAATTTCTACGCGCCAGTATATTTGTTACT 5536
QY 421 TCACCCCTGACATCTCTGTGGCTCTGGCACTTAATTTAGTCTGTATGAGCCCGCCCTT 480
DB 5537 TCACCCCTGACATCTCTGTGGCTCTGGCACTTAATTTAGTCTGTATGAGCCCGCCCTT 5596
QY 481 AATCCGCTCTCGCGTGCAGGACGGTACTAATCTACATTAATGCGCCAGAGGCTCC 540
DB 5597 AATCCGCTCTCGCGTGCAGGACGGTACTAATCTACATTAATGCGCCAGAGGCTCC 5656
QY 541 AATATGACAGTACCGGGTTGCCCGCTACTATCCGTTACCGGCCCTTAGTGCCTAAT 600
DB 5657 AATATGACAGTACCGGGTTGCCCGCTACTATCCGTTACCGGCCCTTAGTGCCTAAT 5716
QY 601 GCAGTTGGAGGCTATGCTATATCAATTTCTTCTGGCTCAAAACAACCCCTACA 660
DB 5717 GCAGTTGGAGGCTATGCTATATCAATTTCTTCTGGCTCAAAACAACCCCTACA 5776
QY 661 TCTGTTGACATGAATTCATTTCTTCTACATGATGTCAGGATTTCTTCAACCTGGCATA 720
DB 5777 TCTGTTGACATGAATTCATTTCTTCTACATGATGTCAGGATTTCTTCAACCTGGCATA 5836

QY 721 GCATCTGAATTGGTTCATCCCAAGCGAGCGCCTTCACTACCGCAATCAAGGTTGGCGCTCG 780
DB 5837 GCATCTGAATTGGTTCATCCCAAGCGAGCGCCTTCACTACCGCAATCAAGGTTGGCGCTCG 5896
QY 781 GTTGAGACATCTGGTGTGTTGCTGAGGAGAAAGCCACTCCGGTCTTGTCATGTTATCAATA 840
DB 5897 GTTGAGACATCTGGTGTGTTGCTGAGGAGAAAGCCACTCCGGTCTTGTCATGTTATGATA 5956
QY 841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCTGGCTTACTG 900
DB 5957 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCTGGCTTACTG 6016
QY 901 GACTTTGCCCTTAGAGCTTGAGTTTGCATCTCACCACTGTAAACCAATACACGTGTG 960
DB 6017 GACTTTGCCCTTAGAGCTTGAGTTTGCATCTCACCACTGTAAACCAATACACGTGTG 6076
QY 961 TCCCGTTACTCCAGCACTGCTCGTCACTCCGCGGAGGCGGACCGGACTGCGGAGCTG 1020
DB 6077 TCCCGTTACTCCAGCACTGCTCGTCACTCCGCGGAGGCGGACCGGACTGCGGAGCTG 6136
QY 1021 ACCCAACTGCGAGCCACCAAGGTTTCATGAAGATCTCCACTTTACCGGCCCTTAATGGGTA 1080
DB 6137 ACCCAACTGCGAGCCACCAAGGTTTCATGAAGATCTCCACTTTACCGGCCCTTAATGGGTA 6196
QY 1081 GGTGAAGTCGCGCGGAGTAGCTCTAACATTAATTAACCTTGCTGACACGCTCTCTCGGC 1140
DB 6197 GGTGAAGTCGCGCGGAGTAGCTCTAACATTAATTAACCTTGCTGACACGCTCTCTCGGC 6256
QY 1141 GGGCTCCCGACAGAAATTAATTTGCTGGCTGGGGGCAACTGTTTATTTCCCGCCCGGTT 1200
DB 6257 GGGCTCCCGACAGAAATTAATTTGCTGGCTGGGGGCAACTGTTTATTTCCCGCCCGGTT 6316
QY 1201 GTCTAGCCCAATGGCGAGCCACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAGCAG 1260
DB 6317 GTCTAGCCCAATGGCGAGCCACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAGCAG 6376
QY 1261 GATAAGGGTGTGCTATCCCGACAGATATCGATCTTGGTGATTTGGGTGATTTGGTCT 1320
DB 6377 GATAAGGGTGTGCTATCCCGACAGATATCGATCTTGGTGATTTGGGTGATTTGGTCT 6436
QY 1321 GATTATGACAAACAGCATGAGCAGGATCGGCCACACCCCGTGGCTGGCGCATCTCGGCT 1380
DB 6437 GATTATGACAAACAGCATGAGCAGGATCGGCCACACCCCGTGGCTGGCGCATCTCGGCT 6496
QY 1381 TTTTCTGTTCTCGAGCAATGATGTTACTTTGGCTGCTCTCCTCCTCAGCCGAGATGAC 1440
DB 6497 TTTTCTGTTCTCGAGCAATGATGTTACTTTGGCTGCTCTCCTCCTCAGCCGAGATGAC 6556
QY 1441 CAGTCCACTTACGGGTCTGCAACTGGCCCGGTTTATATCTCGGACAGCGTGAATTTGGTG 1500
DB 6557 CAGTCCACTTACGGGTCTGCAACTGGCCCGGTTTATATCTCGGACAGCGTGAATTTGGTG 6616
QY 1501 AATGTTGCACTGCGCGCAGCGCGTAGCCGATCGCTTGAATGCTGTCGAAAGTCAACCTC 1560
DB 6617 AATGTTGCACTGCGCGCAGCGCGTAGCCGATCGCTTGAATGCTGTCGAAAGTCAACCTC 6676
QY 1561 GACGCGCGCCCTCCCGACTGTTGAGCAATATTCAGACATTTCTTTGCTGCCCTT 1620
DB 6677 GACGCGCGCCCTCCCGACTGTTGAGCAATATTCAGACATTTCTTTGCTGCCCTT 6736
QY 1621 CGTGCAAGCTCTCTCTTTTGGGAGCGCGCACAAAGAGAGGTTTATCTTATAATAT 1680
DB 6737 CGTGCAAGCTCTCTCTTTTGGGAGCGCGCACAAAGAGAGGTTTATCTTATAATAT 6796
QY 1681 AATACTACTGCTGAGCAGGATCTGATTGAAATGCTGCGGCCATCGGGTCGCCT 1740
DB 6797 AATACTACTGCTGAGCAGGATCTGATTGAAATGCTGCGGCCATCGGGTCGCCT 6856
QY 1741 TCAACCTATACCAACAGGCTTGGGGCGGTCGCGTCCGCTCGCCATTTCTCGCGCGCGGTTTG 1800
DB 6857 TCAACCTATACCAACAGGCTTGGGGCGGTCGCGTCCGCTCGCCATTTCTCGCGCGCGGTTTG 6916
QY 1801 GCTCCACGCTCCGCCCTGGCTCTGCTGGAGGATACATTTTGAATTTATCGGGGCGGCGCAC 1860

[illegible]

RESULT 7

RESOL 7
AAF83496

AAE83496
ID AAF83496 standard: cDNA: 7171 BP.

XX AAF83496;

DT 23-JUL-2001 (first entry)

Mexican strain ET-NANB viral ET1.1 composite cDNA sequence.

KW HEV; enterically transmitted nonA/nonB viral hepatitis agent; viral;

KW ET-NANB; infection; vaccine; ss.

OS Hepatitis E virus.

PN US6229005-B1.

PD 08-MAY-2001.

PF 03-AUG-1998; 98US-00128275.

PR 17-JUN-1988; 88US-00208997.

PR 11-APR-1989; 89US-00336672.

PR 16-JUN-1989; 89US-00367486.

PR 13-OCT-1989; 89US-00420921.

PR 05-APR-1990; 90US-00505888.

PR 05-APR-1991; 91US-00681078.

PR 25-JUL-1994; 94US-00279823.
yy

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (GENE -) GENELABS TECHNOLOGIES INC.
YY

PI Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam AW;

PI Fry KE;
vv

[illegible]

Db 6157 CTCACCAACCGGTGCTACTCGCTTCATGAAGGACCTCTATTTTACTAGTACTAAUGGT 6216
Qy 1078 GTAGTGAAGTCGGCCGGGATAGCTCTAACATTACTTTAACCTTGCTGACACGGCTCCTC 1137
Db 6217 GTTGGTGAGATCGGCCGGGATAGCGCTTACCCCTGTTTAACTTGCTGACACCCCTGCTT 6276
Qy 1138 GCGGGCTCCCGACAGAAATAATTCGTGCGTGGCGGCAACTGTTTTATTTCGCGCCG 1197
Db 6277 GCGGCTTACCGACAGAAATGATTTCTGCTGGCTGGTGGCCAGCTGTTCTACTCTCGCCCC 6336
Qy 1198 GTTGTCTCAGCAATGGCGACCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAG 1257
Db 6337 GTCGTCTCAGCAATGGCGACCGACTGTTAAGCTGTATACATCTGTGGAGAAATGCTCAG 6396
Qy 1258 CAGGATAGGGGTGTGCTATCCGCCACGATATCGATCTTGTGTGATTCGCGTGTGTCATT 1317
Db 6397 CAGGATAGGGGTAATTGCAATCCCGCATGACATCGACCTCGGGGAATCCCGTGTAGTTATT 6456
Qy 1318 CAGGATTAATGACAAACGACATGAGCAGGATCGGCCACCCCGTGGCCCTGCGCCATCTCGG 1377
Db 6457 CAGGATTAATGACAAACCAACATGAGCAGGACCGACCGACACCTTCCCGACGCCATCGCGT 6516
Qy 1378 CCTTTTCTGTCTCCGAGCAAAATGATGTACTTTGGTGTCCCTCACTGACGCGCGAGTAT 1437
Db 6517 CCTTTTCTGTCTCCGAGCTAAAGATGTGCTTTGGCTTCTCTCACCGCTGCCGAGTAT 6576
Qy 1438 GACCAGTCCACTTACGGGTCTCAACTGGCCGGTTTATATCTCGGACAGCGTGACTTTG 1497
Db 6577 GACCAGTCCACTTACGGGTCTTTCGACCGGCCCGAGTCTATGCTCTGACTCTGTGACCTTG 6636
Qy 1498 GTGAATGTTGCGACTGGCGCGCAGGCCGTAGCCCGATCGCTTGACTGGTCCAAAAGTCACC 1557
Db 6637 GTTAATGTTGCGACCGCGCGCAGGCCGTTCGCCGGTCACTCGACTGACCAAGGTACA 6696
Qy 1558 CTCGACGGCGGCCCTCCCGACTGTTGAGCAATATCCAAAGAAATTCCTTTGTGCTCCCC 1617
Db 6697 CTTGATGTCGCCGCCCTTTCACCATCCAGCAGTATTCAAAAGACCTTCTTTGCTCGCG 6756
Qy 1618 CTTGTCGCAAGCTCTCCTTTTGGAGGCCGGCACACAAAAGCAGGTTATCCTTTATAAT 1677
Db 6757 CTCGCGGTAAAGCTCTCCTTTTGGAGGCCAGGAACTACTAAAGCGGGTACCCCTTATAAT 6816
Qy 1678 TATAATACTACTGTAGTGACAGATTCTGATTGAAAATGTCGCGGCCATCGGCTGCC 1737
Db 6817 TATAACACCACTGTAGTGACCAACTGCTCGTTGAGAAATGCGCTGGCATCGGTTGCT 6876
Qy 1738 ATTTCAACCTATACACACAGGCTTGGGCCGGTCCGFTCGCAATTTCTGCGGCCCGCGTT 1797
Db 6877 ATTTCCACCTACACTACTAGCTGGGTGCTGGCCCGGTCTCTAATTTCCGCGGTTGCTGTT 6936
Qy 1798 TTGGCTCAGCTCGGCCCTGGCTCTGCTGGAGGATACTTTTGATATATCCGGGGGGCG 1857
Db 6937 TTAGCCCCCACTCTGTCTAGCATTTGCTTGAGGATACCATGSACTACCCCTGCCCGGCC 6996
Qy 1858 CACACATTTGATGACTTCTGCCCTGAATGCGCGCTTTAGGCCCTCCAGGGTTGTCTTTC 1917
Db 6997 CATACTTTCGATGACTCTGCCCGAGTGCGGCCCTTGGCCCTCCAGGGTTGTCTTTT 7056
Qy 1918 CAGTCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAAGTGGGTAAAACCTCGGAGTTG 1977
Db 7057 CAGTCTACTGTGCTGAGCTTCAGCGCTTAAAGATGAAGTGGGTAAAACCTCGGAGTTA 7116
Qy 1978 TAG 1980
Db 7117 TAG 7119

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 10:14:02 ; Search time 5109.13 Seconds
(without alignments)
11572.836 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096

Perfect score: 1980

Sequence: 1 ATGGCCCTAGGCTCTTTT.....GTAAACTCGGAGTTGTAG 1980

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vit:*

21: em_gss_fun:*

22: em_gss_nam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	62.4	3.2	925	29	CNS0091P	AL053013 Drosophila
c 2	61.8	3.1	1201	13	BX360624	BX360624 BX360624
c 3	60.2	3.0	844	29	CNS0052P	AL056652 Drosophila
c 4	59.4	3.0	1103	13	BX403654	BX403654 BX403654

5	58.2	2.9	1013	29	CNS016KT	AL106871 Drosophila
c 6	57.6	2.9	1000	13	BX407619	BX407619 BX407619
c 7	57.4	2.9	1101	29	CNS017SY	AL108460 Drosophila
c 8	57.2	2.9	1203	29	CNS015Y4	AL106054 Drosophila
9	56.6	2.9	828	13	BX425797	BX425797 BX425797
10	54.8	2.8	1201	13	BX381961	BX381961 BX381961
c 11	54	2.7	1201	9	AL574009	AL574009 AL574009
c 12	53.8	2.7	839	29	CNS004NB	AL054280 Drosophila
c 13	53.8	2.7	1193	13	BX421125	BX421125 BX421125
c 14	53.2	2.7	993	13	BQ144482	BQ144482 NF085H11D
15	53	2.7	1201	13	BX356664	BX356664 BX356664
c 16	52.6	2.7	1144	13	BX415926	BX415926 BX415926
17	52.4	2.6	807	13	BX464554	BX464554 BX464554
c 18	52	2.6	1201	9	AL538546	AL538546 AL538546
c 19	51.8	2.6	1057	13	BX349688	BX349688 BX349688
c 20	51.8	2.6	1201	13	BX335523	BX335523 BX335523
c 21	51.2	2.6	1036	9	AL550172	AL550172 AL550172
c 22	50.8	2.6	1201	29	CNS015WQ	AL106004 Drosophila
c 23	50.6	2.6	877	12	B1956655	B1956655 HVSME000
c 24	49.2	2.5	1009	29	CNS010EW	AL098882 Drosophila
c 25	49	2.5	975	23	CNS02KED	AL201298 Tetradon
c 26	49	2.5	1100	29	CNS001BC	AL060161 Drosophila
c 27	49	2.5	1201	9	AL525224	AL525224 AL525224
c 28	48.8	2.5	935	29	CNS006XK	AL066051 Drosophila
c 29	48.6	2.5	1189	13	BX407485	BX407485 BX407485
c 30	48.4	2.4	863	12	B1953455	B1953455 HVSME001
c 31	48.2	2.4	882	28	BZ050854	BZ050854 jnf64f08.
c 32	48.2	2.4	1061	13	BX391246	BX391246 BX391246
33	48.2	2.4	1071	29	CNS000EM	AL078875 Drosophila
34	48.2	2.4	1101	29	CNS00LXJ	AL078875 Drosophila
c 35	48	2.4	359	14	CB908167	CB908167 trico84XJ
c 36	47.8	2.4	995	29	AG060386	AG060386 Pan trogl
c 37	47.8	2.4	1101	29	CNS016HG	AL106750 Drosophila
c 38	47.8	2.4	1160	29	AG043473	AG043473 Pan trogl
c 39	47.8	2.4	1201	13	BX340485	BX340485 BX340485
c 40	47.6	2.4	689	29	AG044460	AG044460 Pan trogl
c 41	47.6	2.4	906	29	AG081338	AG081338 Pan trogl
c 42	47.6	2.4	1136	28	CC292742	CC292742 CH261-185
c 43	47.4	2.4	1184	13	BX425067	BX425067 BX425067
c 44	47.2	2.4	989	29	AG063827	AG063827 Pan trogl
c 45	47.2	2.4	1273	12	BM562099	BM562099 AGENCOURT

ALIGNMENTS

RESULT 1

CNS0091P

LOCUS

DEFINITION

AL053013

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CNS0091P 925 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RPCL-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

AL053013 GI:4934461

GSS.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 925)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Ooeogawa and

Aaron Mammoser in Pieter de Jong's laboratory in the Department of

TITLE Full-length cDNA libraries and normalization
UNPUBLISHED (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12933798.
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10873.f For
more information about this cluster, see
<http://www.genoscope.cns.fr/>
cgi-bin/cluster.cgi?seq=CSDD1053BA02NP1&cluster=10873.f. Contact :
Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSDD1053BA02NP1.

```

FEATURES
source
1. .1201
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1053YA04"
/tissue.type="PLACENTA"
/clone.lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

[illegible]

RESULT 12	CNS004NB/c	839 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	CNS004NB/c				
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BAC10816 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL054280				
VERSION	AL054280.1	GI:4931788			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	1 (bases 1 to 839)				
TITLE	Genoscope.				
JOURNAL	Direct Submission				
COMMENT	Submitted (02-JUN-1999) Genoscope - Centre National de Sequence ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila				

melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial, EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://bacpac.med.buffalo.edu/drosophila/bac.htm>.

FEATURES
source

ORIGIN

[illegible]

RESULT 13	ACCESSION	REFERENCE
BX421125/c	VERSION	AUTHORS
LOCUS	KEYWORDS	TITLE
DEFINITION	SOURCE	JOURNAL
	ORGANISM	COMMENT

BX4211125 1193 bp mRNA linear EST 15-MAY-2003
 BX4211125 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
 clone CS0DG001YM04 3-PRIME, mRNA sequence.
 BX4211125
 BX4211125.1 GI:30768081
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1193)
 Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: secre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1904.f For

more information about this cluster, see
[http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSODG001BG02NP1&cluster=3904.f](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODG001BG02NP1&cluster=3904.f). Contact :
 Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com> URL : <http://fulllength.invitrogen.com> Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CSODG001BG02NP1

FEATURES
source

1. l1193
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="CS0DG001Y004"
/tissue_type="B CELLS (R
/cell_line="RAMOS CELL L
/clone_lib="Homo sapiens
/note="Vector: pCMVSPORT
with a NotI-oligo(dT) pr
double-strand cDNA was
the Not I and EcoRV site
Library was not normalized

ORIGIN

```

atch      2.7%; Score 53.8; DB 13; Length 1193;
Similarity 37.2%; P-red. No. 0.079;
132; Conservative 41; Mismatches 182; Indels 0; Gaps 0;

48 GTTGGCCGGCGCACACCGGTGACCGGTGCGCGCGTGTGGCGCGCGCAGCGGGCG 107
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
432 GTCGGGCGGGGSCSGSGGCGCCCGGGSGCGGSCCGGGCGCGSCCCCGCGG 373
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

108 TACGGCGGCTGTGTTCTTGGGGTGACCGGGTTGATTCTCAGCCCTTTCGAAATCCCTATAT 167
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
372 GCGGCGCCGGGGCCCGGGCGKGGSGCCNCCNCTTTTCCCTCTTGTCCTCNGCGN 313
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

168 TCATCAACAAACCTTTTGCCCGACAGTTGCGGTGCGGTCTGGAGCTTCGGCT 227
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
312 SNGCCGCTCCYCCCCNTNTBCNNGGGGGCCCGGCGCGGCGCGCGCGCGCGCC 253
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

228 TCGCACAACGCGGGCACTTGGCTTCACATTGGCGAGATCAGGCGCACGGCCCTCCGC 287
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
252 CGCGCTTCGCCCTTCGSCCGCCCGCCCGCCCGSCSCTCTCTCCCGCCCGCCCGCC 193
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

288 TGCGTCCGCTGCGCGACTGCCACAGCGGGGTGCGGGCTGAGCGGTGTGGCGCTTGC 347
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
192 CCCCCCCCCCCCCCCCCCCCCCNVCSGCCCCCCCCCCCCSSSSCSGSCSCSSS 133
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

348 CCATGACACCTCACCCGTCGCCGACGTTGATTCTCGCGGTGCAATCTTAAGCGCG 402
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
132 SCCSSSCSSSSCSGGGGCCCCCTCTGCSGSCSSCSGCTTCTTCBSGCSGCC 78
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 14

BQ144482	BQ144482	993 bp	mRNA	linear	EST 24-APR-2002			
LOCUS	NF085H1D1F1095	Drought	Medicago truncatula	CDNA clone	NF085H1D1			
DEFINITION	5', mRNA sequence.							
ACCESSION	BQ144482							
VERSION	BQ144482.1	GI:20281541						
KEYWORDS	EST.							
SOURCE	Medicago truncatula (barrel medic)							
ORGANISM	Medicago truncatula							
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.							
	1 (bases 1 to 993)							
	Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.							
	Expressed Sequence Tags from the Samuel Roberts Noble Foundation							
TITLE	Medicago truncatula drought library							
JOURNAL	Unpublished (2000)							
COMMENT	Contact: May GD Plant Biology Division							

The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmaw@noble.org
Insert Length: 993 Std Error: 0.00
plate: 085 row: H column: 11
Seq primer: TCACACAGGAAACAGCTATGAC.

FEATURES
SOURCE

```

1993
organism="Medicago truncatula"
mol_type="mRNA"
db_xref="taxon:3880"
clone="NF085H1bT"
tissue_type="Plantlets"
dev_stage="Pooled timepoints"
clone_lib="Drought"
note="Vector: Lambda Zap; Cont
plantlets harvested in a series
timepoints."

```

ORIGIN

Query Match	2.7%;	Score 53.2;	DB 13;	Length 993;
Best Local Similarity	48.7%;	Pred. No. 0.1;		
Matches 145;	Conservative 0;	Mismatches 153;	Indels 0;	Gaps 0;
75	GTCTGGCCCGCTCGTGGCGCGCGACAGCGCGGTACCGCGGTGGTGTCTTGGGGTGAACG	134		
Qy				
79	GGCCGAGGCCGCGGGGGGGCGCGGGGAGGGCCCGCGGGGGCGCGCGCGGG	138		
Db				
135	GGTTGATTCAGACCTTCGCAATCCCTATATTCATCAACCAACCCCTTTGCCCCAGA	194		
Qy				
139	CGGGCCCCCGACACGGCCCCCGCCCCCAAGTTCGCCCGCGCCACCTTTTCCCCCCC	198		
Db				
195	CGTTGCGCTGGGTCCGGGTCTGGACCTCGCTTCGCCCAACACGACCGCGGCACCTTGGGCTC	254		
Qy				
199	CCCTCCCTTCTATCTCGCCCCCCCCCTCCCTCTCTCCCTCCCTCTCTTTCGCCCC	258		
Db				
255	GACTTGGGAGATCAGGCCCGAGCGCCCCCTCGCTGCTCCCTCGCGCACTGCCACAGC	314		
Qy				
259	CCTCCCTCTCCCCCCCCCCCCCCCCCTCAGGCTCTCTTGGCCCGGCTTCGCCCCC	318		
Db				
315	CGGGCTCGGGCTGACGGCTGTGGCGGTGCCCATGACACTCAACCGTCCGGAC	372		
Qy				
319	CGGGCCCCCGCCCCCGGGCCCCCCCCCGCCCCCGCCCCCGCCCCCGCCCCCGTC	376		
Db				

RESULT 15
BX356664

LAC95666	BX356664	1201 bp	mRNA	linear	EST 05-MAY-2003
LOCUS	BX356664	Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
DEFINITION		clone CS0D1015YR03 3-PRIME. mRNA sequence.			

ACCESSION

VERSION	1	GI:30378083
KEYWORDS	EST	
SOURCE	Human	
ORGANISM	Human	
REFERENCE	1	(bases 1 to 1201)
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
TITLE	Full-length cDNA libraries and normalization	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Genoscope	

15771103

BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Catalog 1600
Paradise Avenue Genoscope sequence ID : CS0D1015CA02NP1.

FEATURES

source


```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="CSOD1015YB03"  
/tissue_type="PLACENTA COT 25-NORMALIZED"  
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```
Query Match      2.7%; Score 53; DB 13; Length 1201;  
Best Local Similarity 10.5%; Pred. No. 0.12;  
Matches 43; Conservative 205; Mismatches 161; Indels 0; Gaps 0;  
  
QY      5  GCCTAGGCCTCTTTGCTGTGTTCTCTGTTCTGCTAATGTTGCCCGCCACCGA 64  
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :  
697  SSSSTTSSTTSBSSSTTTTTSBTSSTTSBTSSTTSBTSSTTSBTSSTTSBTS 756  
QY      65  CCGGTACGCGTCTGCGCGCTGCTGGCGCGCAGCGCGGTACCGCGGTGTTCT 124  
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :  
757  TSCTTSSSTBYTBSSSYBBSBBSBTSSCSCTSSBTSSTBSMTSSBCTCTSSSSSS 816  
QY      125  GGGGTGACCGGGTTGATCTCAGCCCTTCGCAATCCCTATATTCACCAACACCCCT 184  
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :  
817  SBTTSTTTSKSGBTTSBTTBSSSSTSSWTBTCTSTTNTTSSYSBSSTSBSTS 876  
QY      185  TTGCCCGACAGTTGCGGCTGCTCCGGGTCTGACCTCGCCTTCGCCAACACCGCCGC 244  
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :  
877  STBSSWSBTTTTSBTSBTBTTSSTTSSTTTBSSSSSSSSSTTTBSSSBTTSSBS 936  
QY      245  CACTTGCTCCACTGCGGAGATCAGCCCGCAGCCCTCCGCTCCCTCCGTCGCCGAC 304  
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :  
937  SSSSSTBSSTSSBSBSSSBTSKSSSSSSSSSSSSSTTBSSTTTTIVTSSBTTSSBT 996  
QY      305  CTGCCACAGCGGGGCTCGCGGCTGACGGCTGTGGCGCCTGCCCATCACACCTCACCCG 364  
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :  
997  TTBSSTTTSSSTBSSBSBSSNTSSSSSSSSSSSTBTSSSNKSSSSSSSSSSSSSS 1056  
QY      365  TCCCGGACGTGATCTCGCGGTGCAATTTCTACCGCGCCAGTAAATTT 413  
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :  
1057  SSSSBTTTBTBBSBTTBSSSSBSTTTTTSBSSSSSTSBTTSSTTTT 1105
```

Search completed: August 21, 2004, 18:05:41
Job time : 5114.13 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 10:24:37 ; Search time 140.394 Seconds
(without alignments)
7826.585 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096

Perfect score: 1980

Sequence: 1 ATGCGCCCTAGGCTCTTTT.....GTAAACTCGGAGTGTTAG 1980

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1980	100.0	1980	1	US-08-240-049B-2
2	1980	100.0	2055	1	US-08-259-148A-4
3	1980	100.0	2055	1	US-08-484-054-4
4	1980	100.0	2055	2	US-07-876-941A-4
5	1980	100.0	2055	3	US-08-477-292-2
6	1980	100.0	2055	4	US-07-870-985A-4
7	1980	100.0	2100	1	US-08-259-148A-2
8	1980	100.0	2100	1	US-08-484-054-2
9	1980	100.0	2100	2	US-07-876-941A-2
10	1980	100.0	2100	4	US-07-870-985A-2
11	1980	100.0	7171	3	US-08-478-507-10
12	1980	100.0	7171	3	US-09-128-275A-10
13	1980	100.0	7171	4	US-09-553-427-10
14	1957.4	98.9	2058	3	US-08-542-634-2
15	1957.4	98.9	2058	5	PCT-US95-13703-2
16	1644	83.0	1644	3	US-08-477-292-4
17	1621.4	81.9	1647	3	US-08-542-634-4
18	1621.4	81.9	1647	5	PCT-US95-13703-4
19	1378.2	69.6	7168	3	US-08-840-316-4
20	1378.2	69.6	7168	3	US-08-809-523-4
21	1378.2	69.6	7168	3	US-08-471-971-4
22	1378.2	69.6	7168	4	US-09-402-776-4
23	1378.2	69.6	7168	5	PCT-US93-08849A-4
24	1378.2	69.6	7168	5	PCT-US93-08849-4
25	1371.8	69.3	1984	4	US-09-172-699-1
26	1368.6	69.1	1983	1	US-08-240-049B-1
27	1368.6	69.1	2049	1	US-08-259-148A-3

28 1368.6 69.1 2049 1 US-08-484-054-3 Sequence 3, Appli
29 1368.6 69.1 2049 2 US-07-876-941A-3 Sequence 3, Appli
30 1368.6 69.1 2049 3 US-08-542-634-1 Sequence 1, Appli
31 1368.6 69.1 2049 3 US-08-477-292-1 Sequence 1, Appli
32 1368.6 69.1 2049 4 US-07-870-985A-3 Sequence 3, Appli
33 1368.6 69.1 2049 5 PCT-US95-13703-1 Sequence 1, Appli
34 1368.6 69.1 2094 1 US-08-259-148A-1 Sequence 1, Appli
35 1368.6 69.1 2094 1 US-08-484-054-1 Sequence 1, Appli
36 1368.6 69.1 2094 2 US-07-876-941A-1 Sequence 1, Appli
37 1368.6 69.1 2094 4 US-07-870-985A-1 Sequence 1, Appli
38 1368.6 69.1 7195 3 US-08-478-507-6 Sequence 6, Appli
39 1368.6 69.1 7195 4 US-09-128-275A-6 Sequence 6, Appli
40 1368.6 69.1 7195 4 US-09-553-427-6 Sequence 6, Appli
41 1308 66.1 1308 1 US-08-259-148A-10 Sequence 10, Appli
42 1308 66.1 1308 1 US-08-484-054-10 Sequence 10, Appli
43 1308 66.1 1308 2 US-07-876-941A-10 Sequence 10, Appli
44 1308 66.1 1308 4 US-07-870-985A-10 Sequence 10, Appli
45 1289.2 65.1 7207 4 US-09-462-606-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-240-049B-2
; Sequence 2, Application US/08240049B
; Patent No. 5686239
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Tam, Albert W.
; APPLICANT: Varbough, Patrice O.
; TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,049B
; FILING DATE: 09-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles K. Sholtz
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.35; G32P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1980 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico Strain)
; INDIVIDUAL ISOLATE: ORF-2 region
US-08-240-049B-2

Query Match 100.0%; Score 1980; DB 1; Length 1980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1321 GATTATGACAAACAGCATGAGCAGGATCGGCCCAACCCCGTGCCTCGCCGATCTCGGCT 1380
1321 GATTATGACAAACAGCATGAGCAGGATCGGCCCAACCCCGTGCCTCGCCGATCTCGGCT 1380
1381 TTTTCTGTTCTCCGAGCAATGATGATCTTGGTGTCCCTCACTGACCGAGTATGAC 1440
1381 TTTTCTGTTCTCCGAGCAATGATGATCTTGGTGTCCCTCACTGACCGAGTATGAC 1440
1441 CAGTCACCTAGGGTCTCACTGGCCCGGTTTATATCTCGACAGCGTGTGTTGG 1500
1441 CAGTCACCTAGGGTCTCACTGGCCCGGTTTATATCTCGACAGCGTGTGTTGG 1500
1501 AATGTTGCGACTGGCGCAGCGCGTAGCCGATCGCTTGAATGCTGCTCAAGTCAACCTC 1560
1501 AATGTTGCGACTGGCGCAGCGCGTAGCCGATCGCTTGAATGCTGCTCAAGTCAACCTC 1560
1561 GACGGCGCGCCCTCCCGACTGTTGAGCAATATTCAGACATTTCTTGTGCTCCCGCTT 1620
1561 GACGGCGCGCCCTCCCGACTGTTGAGCAATATTCAGACATTTCTTGTGCTCCCGCTT 1620
1621 CGTGGCAAGCTCTCTTTTGGGAGCGCGGCACAAAGAGGTTATCTTATATAT 1680
1621 CGTGGCAAGCTCTCTTTTGGGAGCGCGGCACAAAGAGGTTATCTTATATAT 1680
1681 AATACTACTGCTAGTGACAGATTTCTGATTTGAAATGCTGCGGCCATCGGCTCGCAT 1740
1681 AATACTACTGCTAGTGACAGATTTCTGATTTGAAATGCTGCGGCCATCGGCTCGCAT 1740
1741 TCAACCTATACCAACAGGCTTGGGCGGCTCGGTCGCAATTTCTGCGCGCGGTTTG 1800
1741 TCAACCTATACCAACAGGCTTGGGCGGCTCGGTCGCAATTTCTGCGCGCGGTTTG 1800
1801 GCTCCAGCTCCCGCTGCTGCTGAGGATCTTTGATTTATCCGGCGCGGCGCAC 1860
1801 GCTCCAGCTCCCGCTGCTGCTGAGGATCTTTGATTTATCCGGCGCGGCGCAC 1860
1861 ACATTTGATGACTTCTGCGCTGAATCCGCGCTTTAGGCTCCAGGCTTGTCTTCCAG 1920
1861 ACATTTGATGACTTCTGCGCTGAATCCGCGCTTTAGGCTCCAGGCTTGTCTTCCAG 1920
1921 TCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAAACTCGGGAGTTGTAG 1980
1921 TCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAAACTCGGGAGTTGTAG 1980

RESULT 3

US-08-484-054-4
; Sequence 4, Application US/08484054
; Patent No. 5770689
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krwaczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,054
; FILING DATE: 07-JUN-1995

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
FILING DATE: 20-APRIL-1992
PRIOR APPLICATION DATA:
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Shultz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.38
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2055 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 7
US-08-484-054-4

Query Match 100.0%; Score 1980; DB 1; Length 2055;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTAGGCTCTTTGCTGTTGTTCTCTTGTTCCTGCTATGTTGCCGCGCCA 60
DB 1 ATGCGCCCTAGGCTCTTTGCTGTTGTTCTCTTGTTCCTGCTATGTTGCCGCGCCA 60
QY 61 CCGACCGGTACGCGCTCTGCGCCCGCTGCTGGCGCGGCGGCGGTACCGGCGGTGGT 120
DB 61 CCGACCGGTACGCGCTCTGCGCCCGCTGCTGGCGCGGCGGCGGTACCGGCGGTGGT 120
QY 121 TTCTGGGGTACCGGGTGTGATCTCAGCCCTTCCGATCCCTATATTCATCCAAACCAAC 180
DB 121 TTCTGGGGTACCGGGTGTGATCTCAGCCCTTCCGATCCCTATATTCATCCAAACCAAC 180
QY 181 CCCTTTGCCCGACAGCTTGGCGCTGCTGCGGCTCTGGACCTCGCCCTTCGCCCAACAGCC 240
DB 181 CCCTTTGCCCGACAGCTTGGCGCTGCTGCGGCTCTGGACCTCGCCCTTCGCCCAACAGCC 240
QY 241 CGGCCACTTGGCTCCACTTGGCGAGATCAGGCCAGCGCCCTTCGCGTCCCTCCCGTCGC 300
DB 241 CGGCCACTTGGCTCCACTTGGCGAGATCAGGCCAGCGCCCTTCGCGTCCCTCCCGTCGC 300
QY 301 CGACCTGCCACAGCGGGGTGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 360
DB 301 CGACCTGCCACAGCGGGGTGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 360
QY 361 CCGCTCCCGGACGTTGATTCTCGCGGTGCAATTTCTACGCCCGCAGTATAATTTGTCTACT 420
DB 361 CCGCTCCCGGACGTTGATTCTCGCGGTGCAATTTCTACGCCCGCAGTATAATTTGTCTACT 420


```

1141 GGGCTCCGACAGAAATATTCGTCGGCTGGGGGCACTGTTTATTCGCCCGCGTT 1200
1201 GTCTCAGCAATGGGAGCAACCGGTGAAGTCTATACATCATGTGAGAAATGCTCAGCAG 1260
1201 GTCTCAGCAATGGGAGCAACCGGTGAAGTCTATACATCATGTGAGAAATGCTCAGCAG 1260
1261 GATAGGGGTGTTGCTATCCCCACAGATATCGATCTGTTGGTGAATTCGGTGTGCTCATTCAG 1320
1261 GATAGGGGTGTTGCTATCCCCACAGATATCGATCTGTTGGTGAATTCGGTGTGCTCATTCAG 1320
1321 GATTATGACAAACAGCATGAGCAGGATCGGCCACCCCGCTCGCTCGCCATCTCGCCCT 1380
1321 GATTATGACAAACAGCATGAGCAGGATCGGCCACCCCGCTCGCTCGCCATCTCGCCCT 1380
1381 TTTTCTGTTCTCGAGCAAAATGATGATCTTTGGTGTCTCTCACTGACGCGAGTATGAC 1440
1381 TTTTCTGTTCTCGAGCAAAATGATGATCTTTGGTGTCTCTCACTGACGCGAGTATGAC 1440
1441 CAGTCCACTTACGGGTGCTCAACTGGCCCGGTTTATCTCGACAGCGTGATTTGGTG 1500
1441 CAGTCCACTTACGGGTGCTCAACTGGCCCGGTTTATCTCGACAGCGTGATTTGGTG 1500
1501 AATGTTGCACTGGCGCAGGCGGTAGCCCGATCGCTTGAATGCTGCTGCTGCTGCTGCTG 1560
1501 AATGTTGCACTGGCGCAGGCGGTAGCCCGATCGCTTGAATGCTGCTGCTGCTGCTGCTG 1560
1561 GACGGCGGCGCTCCCGTCTGTTGAGCAATATTCAGAGCAATCTTGTGCTCCCGCTT 1620
1561 GACGGCGGCGCTCCCGTCTGTTGAGCAATATTCAGAGCAATCTTGTGCTCCCGCTT 1620
1621 CGTGGCAAGCTCTCTTTTGGGAGCGCGGCACAAACAAAGCAGTTATCTTATATAT 1680
1621 CGTGGCAAGCTCTCTTTTGGGAGCGCGGCACAAACAAAGCAGTTATCTTATATAT 1680
1681 AATACTACTGCTAGTGACAGATTTCTGATTGAAATGCTGCGCGCATCGGGTTCGCATT 1740
1681 AATACTACTGCTAGTGACAGATTTCTGATTGAAATGCTGCGCGCATCGGGTTCGCATT 1740
1741 TCACTCTATACACAGGCTTGGGCGCGGTTCGGTTCGCATCTTCTCGCGCGGGTTTG 1800
1741 TCACTCTATACACAGGCTTGGGCGCGGTTCGGTTCGCATCTTCTCGCGCGGGTTTG 1800
1801 GCTCCAGCTCGCCCTGCTGCTGAGGATACATTTTGAATGCTGCGGGCGGGCGCAC 1860
1801 GCTCCAGCTCGCCCTGCTGCTGAGGATACATTTTGAATGCTGCGGGCGGGCGCAC 1860
1861 ACATTTGATGACTTTCCTGCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
1861 ACATTTGATGACTTTCCTGCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
1921 TCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
1921 TCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980

```

RESULT 6

```

US-07-870-985A-4
; Sequence 4, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto

```

```

; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2055 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 7
; US-07-870-985A-4

```

```

Query Match 100.0%; Score 1980; DB 4; Length 2055;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTAGCGCTCTTTTGTGTTGTTTCCTCTTGTCTTGTCTATGTTGCCGCGCCA 60
Db 1 ATGCGCCCTAGCGCTCTTTTGTGTTGTTTCCTCTTGTCTTGTCTATGTTGCCGCGCCA 60

QY 61 CCGACCGGTGACCGGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CCGACCGGTGACCGGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120

QY 121 TTCTGGGGTGACCGGCTTGAATTCCTCAGCCCTTTCGCAATCCCTTATATTCATCCAAAC 180
Db 121 TTCTGGGGTGACCGGCTTGAATTCCTCAGCCCTTTCGCAATCCCTTATATTCATCCAAAC 180

QY 181 CCCTTTGCCCGCAGAGTTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 240
Db 181 CCCTTTGCCCGCAGAGTTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 240

QY 241 CGGCCACTTTGGCTCCACTTTGGCGAGATCAGGGCCAGCGCCCTCCCGCTCCCGTCGC 300
Db 241 CGGCCACTTTGGCTCCACTTTGGCGAGATCAGGGCCAGCGCCCTCCCGCTCCCGTCGC 300

```

```
QY 301 CGACCTGCCACAGCGGGGCTCGCGCTGACGGCTGTGGGGCTCGCCCATGACACCTCA 360
Db 301 CGACCTGCCACAGCGGGGCTCGCGCTGACGGCTGTGGGGCTCGCCCATGACACCTCA 360
QY 361 CCCGTCGCCGAGTGTGATTCTCGCGGTGCAATTCCTACGCGCGCAGTATAAATTTGCTACT 420
Db 361 CCCGTCGCCGAGTGTGATTCTCGCGGTGCAATTCCTACGCGCGCAGTATAAATTTGCTACT 420
QY 421 TCACCCCTGACATCCTCTGTGGCTCTGGCACTAATTTAGTCCTGTATGCGACCCCTT 480
Db 421 TCACCCCTGACATCCTCTGTGGCTCTGGCACTAATTTAGTCCTGTATGCGACCCCTT 480
QY 481 AATCGCCTCTGCGCCTGCAGGACGGTACTAATACTCACATTAATGGCCACAGAGCCCTCC 540
Db 481 AATCGCCTCTGCGCCTGCAGGACGGTACTAATACTCACATTAATGGCCACAGAGCCCTCC 540
QY 541 AATATGACAGTACCGGGTTGCCGCGCTACTATTCGGTTACCGGCCCTTAGTGCTTAAT 600
Db 541 AATATGACAGTACCGGGTTGCCGCGCTACTATTCGGTTACCGGCCCTTAGTGCTTAAT 600
QY 601 GCAGTTGAGGCTATGCTATATCAATTTCTTCTGGCCTCAAAACAACCAACCCCTACA 660
Db 601 GCAGTTGAGGCTATGCTATATCAATTTCTTCTGGCCTCAAAACAACCAACCCCTACA 660
QY 661 TCTGTTGACATGAATTCATTTCCACTGATGTGAGGATTTCTGTTCAACCTGGCATA 720
Db 661 TCTGTTGACATGAATTCATTTCCACTGATGTGAGGATTTCTGTTCAACCTGGCATA 720
QY 721 GCATCTGAATTTGCTCATCCCAAGCAGCGCTTACTACCGGAATCAAGTTGGCGCTCG 780
Db 721 GCATCTGAATTTGCTCATCCCAAGCAGCGCTTACTACCGGAATCAAGTTGGCGCTCG 780
QY 781 GTTGAGACATCTGGTGTGCTGAGGAGAGCCACTCCGGTCTGTGATGTATGCATA 840
Db 781 GTTGAGACATCTGGTGTGCTGAGGAGAGCCACTCCGGTCTGTGATGTATGCATA 840
QY 841 CATGGCTCCAGTTAACTCCTTACCAATACCCCTTATACGGGTGCCCTTGGCTTACTG 900
Db 841 CATGGCTCCAGTTAACTCCTTACCAATACCCCTTATACGGGTGCCCTTGGCTTACTG 900
QY 901 GACTTTGCTTAGAGTTGAGTTTGGCAATCTCACACCTGTAAACCAATACAGTGTG 960
Db 901 GACTTTGCTTAGAGTTTGGCAATCTCACACCTGTAAACCAATACAGTGTG 960
QY 961 TCCGTTTACTCCAGCACTGCTGCTCACTCCGCGGAGGGCGGAGCTCGCGAGCTG 1020
Db 961 TCCGTTTACTCCAGCACTGCTGCTCACTCCGCGGAGGGCGGAGCTCGCGAGCTG 1020
QY 1021 ACCACAACTGCAGCCACAGGTTTCATGAAAGATCTCCACTTTACCGGCTTAAATGGGTA 1080
Db 1021 ACCACAACTGCAGCCACAGGTTTCATGAAAGATCTCCACTTTACCGGCTTAAATGGGTA 1080
QY 1081 GGTGAAGTGGCGGGGATAGCTCTAAACATTAACCTTGTGACACGCTCTCGGC 1140
Db 1081 GGTGAAGTGGCGGGGATAGCTCTAAACATTAACCTTGTGACACGCTCTCGGC 1140
QY 1141 GGGCTCCGACAGAAATTAATTTGCTGCGCTGCGCGCAACTGTTTATTCGCGCCGGT 1200
Db 1141 GGGCTCCGACAGAAATTAATTTGCTGCGCTGCGCGCAACTGTTTATTCGCGCCGGT 1200
QY 1201 GTCTCAGCAATGGCGAGCCCAACCGTGAAGCTCTATACATCAGTGGAGAATGCTCAGAG 1260
Db 1201 GTCTCAGCAATGGCGAGCCCAACCGTGAAGCTCTATACATCAGTGGAGAATGCTCAGAG 1260
QY 1261 GATTAAGGTGTGTATCCCCACAGATATCGATTTGTGTGATTCGCGTGTGTCATTACAG 1320
Db 1261 GATTAAGGTGTGTATCCCCACAGATATCGATTTGTGTGATTCGCGTGTGTCATTACAG 1320
QY 1321 GATTATGACACCAAGCATGAGCAGATCGGCCACCCCGTCCGCTGCGCATCTCGGCT 1380
Db 1321 GATTATGACACCAAGCATGAGCAGATCGGCCACCCCGTCCGCTGCGCATCTCGGCT 1380
QY 1381 TTTTCTGTCTCCGAGCAATGATGTACTTTTGGCTGTCCCTCACTGACCGCGAGTATGAC 1440
```

```
Db 1381 TTTTCTGTCTCCGAGCAATGATGTACTTTTGGCTGTCCCTCACTGACCGCGAGTATGAC 1440
QY 1441 CAGTCCACTTACGGGTGCTCAACTGGCCGGTTTATATCTCGGACAGCGTGACTTTGGTG 1500
Db 1441 CAGTCCACTTACGGGTGCTCAACTGGCCGGTTTATATCTCGGACAGCGTGACTTTGGTG 1500
QY 1501 AATGTTGCGACTGGCGCGCAGCGCTAGCCGATCGCTTGAATTCGCGGTCGCGGTCACCTC 1560
Db 1501 AATGTTGCGACTGGCGCGCAGCGCGTAGCCGATCGCTTGAATTCGCGGTCGCGGTCACCTC 1560
QY 1561 GACGGCGGGCCCTCCCGACTGTTGAGCAATATTCGAAGACATTTCTTGTGTCTCCCCCTT 1620
Db 1561 GACGGCGGGCCCTCCCGACTGTTGAGCAATATTCGAAGACATTTCTTGTGTCTCCCCCTT 1620
QY 1621 CGTGCAAGCTCTCCTTTTGGGAGCGCGGCAACAACAAAGCAGGTTATCTTATAAATTAAT 1680
Db 1621 CGTGCAAGCTCTCCTTTTGGGAGCGCGGCAACAACAAAGCAGGTTATCTTATAAATTAAT 1680
QY 1681 AATACTACTGCTAGTGACCAATTTCTGATTGCAAAATGCTCGGCGCATCGGGTCGCCAAT 1740
Db 1681 AATACTACTGCTAGTGACCAATTTCTGATTGCAAAATGCTCGGCGCATCGGGTCGCCAAT 1740
QY 1741 TCAACCTATACACACAGGCTTTGGGCGCGGTCGCGCTCGCCATTTCTCGCGCGCGGTTTGG 1800
Db 1741 TCAACCTATACACACAGGCTTTGGGCGCGGTCGCGCTCGCCATTTCTCGCGCGCGGTTTGG 1800
QY 1801 GCTCAGCTCCGCGCTGCTGCTGAGAGATATTTTGAATATTCGCGGCGCGGCGCAC 1860
Db 1801 GCTCAGCTCCGCGCTGCTGCTGAGAGATATTTTGAATATTCGCGGCGCGGCGCAC 1860
QY 1861 ACATTTGATGACTTCTGCCCTGAATGCGCGCTTAGCCCTCCAGGTTGTCTTCCAG 1920
Db 1861 ACATTTGATGACTTCTGCCCTGAATGCGCGCTTAGCCCTCCAGGTTGTCTTCCAG 1920
QY 1921 TCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAAGTGGGTAAAGTGGGTGAG 1980
Db 1921 TCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAAGTGGGTGAG 1980
```

RESULT 7

```
US-08-259-148A-2
; Sequence 2, Application US/08259148A
; Patent No. 5741490
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; APPLICANT: Yarbough, Patrice D.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/259,148A
; FILING DATE: 13-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
```


QY 721 GCATCTGAATTGGTTCATCCCAAGCGAGCGCCTTCACTACCGCAATCAAGGTTGGCGCTCG 780
Db 766 GCATCTGAATTGGTTCATCCCAAGCGAGCGCCTTCACTACCGCAATCAAGGTTGGCGCTCG 825
QY 781 GTTGAGACATCTGGTGTGCTGAGGAGGAAGCACCCTCGGFCCTTGTGCATGTTATGCATA 840
Db 826 GTTGAGACATCTGGTGTGCTGAGGAGGAAGCACCCTCGGFCCTTGTGCATGTTATGCATA 885
QY 841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCCCTTGGCTTACTG 900
Db 886 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCCCTTGGCTTACTG 945
QY 901 GACTTTGCTTGAAGCTTCAGTTTGCATCTCAACACCTGTAAACCAATACAGCTGTG 960
Db 946 GACTTTGCTTGAAGCTTCAGTTTGCATCTCAACACCTGTAAACCAATACAGCTGTG 1005
QY 961 TCCCGTTACTCCAGCACTGCTGTCACCTCGGCCGAGGGGCCGAGCGGACTCGGAGCTG 1020
Db 1006 TCCCGTTACTCCAGCACTGCTGTCACCTCGGCCGAGGGGCCGAGCGGACTCGGAGCTG 1065
QY 1021 ACCAAGCTGACGACCAAGCTTCATGAAGATCTCACTTTACCGGCTTAAATGGGTA 1080
Db 1066 ACCAAGCTGACGACCAAGCTTCATGAAGATCTCACTTTACCGGCTTAAATGGGTA 1125
QY 1081 GGTGAAGTCGGCGCGGATAGCTCTAACATTAATTAACCTTGTGACACGCTCTCTCGGC 1140
Db 1126 GGTGAAGTCGGCGCGGATAGCTCTAACATTAATTAACCTTGTGACACGCTCTCTCGGC 1185
QY 1141 GGGCTCCGACAGAAATTAATTCGTGCGTGGCGGCAACTGTGTTTATTCGCGCCCGGTT 1200
Db 1186 GGGCTCCGACAGAAATTAATTCGTGCGTGGCGGCAACTGTGTTTATTCGCGCCCGGTT 1245
QY 1201 GTCTGAGCAATGGCGAGCCCAAGCTCTATACATCAGTGAGAGATGCTCAGCAG 1260
Db 1246 GTCTGAGCAATGGCGAGCCCAAGCTCTATACATCAGTGAGAGATGCTCAGCAG 1305
QY 1261 GATAAGGCTGCTCTATCCCCACGATATCGATCTTGGTGATTCGGGTGTGTCATTCAG 1320
Db 1306 GATAAGGCTGCTCTATCCCCACGATATCGATCTTGGTGATTCGGGTGTGTCATTCAG 1365
QY 1321 GATTATGACAAACAGCATGAGCAGGATCGGCCCAACCCCGTGGCTCGGCCATCTCGGCT 1380
Db 1366 GATTATGACAAACAGCATGAGCAGGATCGGCCCAACCCCGTGGCTCGGCCATCTCGGCT 1425
QY 1381 TTTTCTGTTCTCGAGCAATGATGATCTTGGCTGTCTCTACTCAGCCGAGTATGAC 1440
Db 1426 TTTTCTGTTCTCGAGCAATGATGATCTTGGCTGTCTCTACTCAGCCGAGTATGAC 1485
QY 1441 CAGTCCACTTACGGGTGCTCAACTGGCCCGGTTTATATCTCGACAGCGTGACTTTGGTG 1500
Db 1486 CAGTCCACTTACGGGTGCTCAACTGGCCCGGTTTATATCTCGACAGCGTGACTTTGGTG 1545
QY 1501 AATGTTGCGACTGGCGGCGAGCGGTAGCCGATCGGTTGACTGTTGCTCAAGTCAAGCTC 1560
Db 1546 AATGTTGCGACTGGCGGCGAGCGGTAGCCGATCGGTTGACTGTTGCTCAAGTCAAGCTC 1605
QY 1561 GAGCGGGCCCTCCCGACTGTTGAGCAATATTCGAAGACATCTTTGTCCTCCCTT 1620
Db 1606 GAGCGGGCCCTCCCGACTGTTGAGCAATATTCGAAGACATCTTTGTCCTCCCTT 1665
QY 1621 CGTGGCAAGCTCTCTTTTGGGAGCGCGGACAAACAAAGCAGGTTATCTTATAATTAT 1680
Db 1666 CGTGGCAAGCTCTCTTTTGGGAGCGCGGACAAACAAAGCAGGTTATCTTATAATTAT 1725
QY 1681 AATACTACTCTAGTGACAGATTCTGATTTGAAATGCTCGCGCCATCGGGTCGCCATT 1740
Db 1726 AATACTACTCTAGTGACAGATTCTGATTTGAAATGCTCGCGCCATCGGGTCGCCATT 1785
QY 1741 TCAACCTATACCAAGCTTTGGGCGGCTCGGTCGCCATTTCTGGCGCGGGTTTTG 1800
Db 1786 TCAACCTATACCAAGCTTTGGGCGGCTCGGTCGCCATTTCTGGCGCGGGTTTTG 1845
QY 1801 GCTCCACGCTCCGCGCTGGCTCTGCTGGAAGATACCTTTTGATTATCCGGGGCGGCGCAC 1860

Db 1846 GCTCCACGCTCGGCCCTGCTCTGCTGGAGGATACCTTTTGATTATCCGGGGCGGCGCAC 1905
QY 1861 ACATTTGATGACTTCTGCGCTGAATGCGCGCTTTAGGCTCCAGGGTTGTGCTTTCCAG 1920
Db 1906 ACATTTGATGACTTCTGCGCTGAATGCGCGCTTTAGGCTCCAGGGTTGTGCTTTCCAG 1965
QY 1921 TCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAAGTTAGGTTGGGTAAGCTCGGAGTTGTAG 1980
Db 1966 TCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAAGTTAGGTTGGGTAAGCTCGGAGTTGTAG 2025
RESULT 9
US-07-876-941A-2
; Sequence 2, Application US/07876941A
; Patent No.: 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; TITLE OF INVENTION: Antibodies
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/876,941A
; FILING DATE: 01-MAY-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.33
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO

ORIGINAL SOURCE:						
; INDIVIDUAL ISOLATE: MEXICO, FIGURE 7						
US-07-876-941A-2						
Query Match 100.0%; Score 1980; DB 2; Length 2100;						
Best Local Similarity 100.0%; Pred. No. 0;						
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	1	ATGCGCCTAGCCCTCTTTTGTCTTGTTTCCTCTTGTTCCTCTTGTTCGCCTATGTTGTCGCCGGCCA	60			
Db	46	ATGCGCCTAGCGCTCTTTTGTCTGTGTTCCTCTTGTTCCTCTTGTTCGCCTATGTTGTCGCCGGCCA	105			
Qy	61	CCGACCGGTGACCGCTGTGCGCGCGTCTGTGGCGGCGCAGCGCGGTGACCGCGGTGGT	120			
Db	106	CCGACCGGTGACCGCTGTGCGCGCGTCTGTGGCGGCGCAGCGCGGTGACCGCGGTGGT	165			
Qy	121	TTCGCGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTTCTATCCAAACCAAC	180			
Db	166	TTCTGGGTTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTTCTATCCAAACCAAC	225			
Qy	181	CCCTTTGCCCCCAGACGTTGCGCGCTGCGTCCGGGTCTGGACCTCGCCCTTCGCAACCAAGCC	240			
Db	226	CCCTTTGCCCCCAGACGTTGCGCGCTGCGTCCGGGTCTGGACCTCGCCCTTCGCAACCAAGCC	285			
Qy	241	CGGCCACTTGGCTCCACTTGGCGAGATCAGGCCCAGCGCCCCCTCCGCTGCCTCCCGTCGC	300			
Db	286	CGGCCACTTGGCTCCACTTGGCGAGATCAGGCCCAGCGCCCCCTCCGCTGCCTCCCGTCGC	345			
Qy	301	CGACCTGCCACAGCGGGGCTGCGCGCTGCAGCGCTGGCGCTGCCCATGACACCTCA	360			
Db	346	CGACCTGCCACAGCGGGGCTGCGCGCTGCAGCGCTGGCGCTGCCCATGACACCTCA	405			
Qy	361	CCCGTCCGGACGTTGATTCTCGCGGTGCAATTTACGCGGCCAGTATAATTTGCTCTACT	420			
Db	406	CCCGTCCGGACGTTGATTCTCGCGGTGCAATTTACGCGGCCAGTATAATTTGCTCTACT	465			
Qy	421	TCACCCCTGACATCCTCTGTGGCTCTGGGCACTAATTTAGTCTGTATGACAGCCCCCCTT	480			
Db	466	TCACCCCTGACATCCTCTGTGGCTCTGGGCACTAATTTAGTCTGTATGACAGCCCCCCTT	525			
Qy	481	AATCCGCTCTGCGCTGCGAGCGGTACTATTAACAATTTAGTCTGTATGACAGCCCCCCTT	540			
Db	526	AATCCGCTCTGCGCTGCGAGCGGTACTATTAACAATTTAGTCTGTATGACAGCCCCCCTT	585			
Qy	541	AATTATGCACAGTACCGGGTTGCCCGGCTACTATCGGTTACC GGCCCTTAGTGCCTAAT	600			
Db	586	AATTATGCACAGTACCGGGTTGCCCGGCTACTATCGGTTACC GGCCCTTAGTGCCTAAT	645			
Qy	601	GCAAGTTGGAGGCTATGCTATATCAATTTCTTTTGGGCTCAAACAACCAACCCCTTACA	660			
Db	646	GCAAGTTGGAGGCTATGCTATATCAATTTCTTTTGGGCTCAAACAACCAACCCCTTACA	705			
Qy	661	TCGTGTCACATGAATCCATTAATTCACATGATGTCAGGATTCCTGTTTCAACCTGGCATA	720			
Db	706	TCGTGTCACATGAATCCATTAATTCACATGATGTCAGGATTCCTGTTTCAACCTGGCATA	765			
Qy	721	GCATCTGAATGGTCAATCCCAAGCGAGCGCTTCACATACCGCAATCAAGGTTGGCGCTCG	780			
Db	766	GCATCTGAATGGTCAATCCCAAGCGAGCGCTTCACATACCGCAATCAAGGTTGGCGCTCG	825			
Qy	781	GTTGAGACATCTGGTGTGTGTGAGGAGGAGCCACCTCCGGTCTGTGTCATGTTATGCATA	840			
Db	826	GTTGAGACATCTGGTGTGTGTGAGGAGGAGCCACCTCCGGTCTGTGTCATGTTATGCATA	885			
Qy	841	CATGGCTCTCCAGTTAACTCGTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG	900			
Db	886	CATGGCTCTCCAGTTAACTCGTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG	945			
Qy	901	GACTTTGCCCTTAGAGCTTGAGTTTCGCAATCTCACCACTGTAAACCAATACACGTTG	960			
Db	946	GACTTTGCCCTTAGAGCTTGAGTTTCGCAATCTCACCACTGTAAACCAATACACGTTG	1005			
Qy	961	TCCGGTTACTCCAGCACTGCTCGTCACTCCGCGCCGAGGGGCCGACGGGACTGCGAGGCTG	1020			

RESULT 10
US-07-870-985A-2
; Sequence 2, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:

QY 301 CGACCTGCCACAGCGGGCTGCGGCTGACGGCTGTGGCCCTGCCCATGACACCTCA 360
DB 5417 CGACCTGCCACAGCGGGCTGCGGCTGACGGCTGTGGCCCTGCCCATGACACCTCA 5476
QY 361 CCCGTCGCCGAGCTGATCTCGCGGTGCAATCTTACGCGCCAGTATAATTTGTACT 420
DB 5477 CCCGTCGCCGAGCTGATCTCGCGGTGCAATCTTACGCGCCAGTATAATTTGTACT 5536
QY 421 TCACCCCTGACATCTCTGTGGCCCTTGGCACTAATTTAGTCTGTATGACAGCCCTT 480
DB 5537 TCACCCCTGACATCTCTGTGGCCCTTGGCACTAATTTAGTCTGTATGACAGCCCTT 5596
QY 481 AATCGCCCTCTCCGCTGCGAGCAGGTACTAATCTCAATATATGCGCCACAGAGCCCTCC 540
DB 5597 AATCGCCCTCTCCGCTGCGAGCAGGTACTAATCTCAATATATGCGCCACAGAGCCCTCC 5656
QY 541 AATATGACACAGTACCGGGTTCGCCGCTACTATCCGTTACCGGCTTACGCTTAAAT 600
DB 5657 AATATGACACAGTACCGGGTTCGCCGCTACTATCCGTTACCGGCTTACGCTTAAAT 5716
QY 601 GCAGTTGAGGCTATGCTATATCCATTTCTTCTGSCCTCAACACACACACCCCTACA 660
DB 5717 GCAGTTGAGGCTATGCTATATCCATTTCTTCTGSCCTCAACACACACACCCCTACA 5776
QY 661 TCTGTTGACATGAATTCATTAATCTTCACTGATGTCAGGATCTTGTTCACCTGGCATA 720
DB 5777 TCTGTTGACATGAATTCATTAATCTTCACTGATGTCAGGATCTTGTTCACCTGGCATA 5836
QY 721 GCATCTGAATGGTCAATCCAAAGCAGCGCTTCACTACCGCAATCAAGGTTGGGCTCG 780
DB 5837 GCATCTGAATGGTCAATCCAAAGCAGCGCTTCACTACCGCAATCAAGGTTGGGCTCG 5896
QY 781 GTTGAGACATCTGGTGTCTGAGGAGGAAGCACTCCGCTTGTCAATGTTATGATA 840
DB 5897 GTTGAGACATCTGGTGTCTGAGGAGGAAGCACTCCGCTTGTCAATGTTATGATA 5956
QY 841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCTTGGCTACTG 900
DB 5957 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCTTGGCTACTG 6016
QY 901 GACTTTGCTTAGAGTTGAGTTTGGCAATCTCAACACCTGTAAACCAATACAGTGTG 960
DB 6017 GACTTTGCTTAGAGTTGAGTTTGGCAATCTCAACACCTGTAAACCAATACAGTGTG 6076
QY 961 TCCGTTACTCCAGCACTGCTGTCTACCTCGCGGAGGCGCAGCGGACTCGGAGCTG 1020
DB 6077 TCCGTTACTCCAGCACTGCTGTCTACCTCGCGGAGGCGCAGCGGACTCGGAGCTG 6136
QY 1021 ACCAACTGCGAGCCACCAAGTTTCATGAAGATCTCCACTTTACCGGCTTAAATGGGTA 1080
DB 6137 ACCAACTGCGAGCCACCAAGTTTCATGAAGATCTCCACTTTACCGGCTTAAATGGGTA 6196
QY 1081 GGTGAAGTGGCGCGGGATAGTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
DB 6197 GGTGAAGTGGCGCGGGATAGTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 6256
QY 1141 GGGCTCCCGACAGAAATTAATTTGCTGCTGGCGGCAACTGTTTATTTCCCGCCGGTT 1200
DB 6257 GGGCTCCCGACAGAAATTAATTTGCTGCTGGCGGCAACTGTTTATTTCCCGCCGGTT 6316
QY 1201 GTCTCAGCCAAATGGCGAGCAACCGTGAAGCTCTATACATCAGTGGAGAAATCTCAGCAG 1260
DB 6317 GTCTCAGCCAAATGGCGAGCAACCGTGAAGCTCTATACATCAGTGGAGAAATCTCAGCAG 6376
QY 1261 GATAGGGTGTGCTATATCCCGACGATATCGATCTTGGTGATTCGGTGTGTCTTATCAG 1320
DB 6377 GATAGGGTGTGCTATATCCCGACGATATCGATCTTGGTGATTCGGTGTGTCTTATCAG 6436
QY 1321 GATTATGACAAACAGCATGAGCAGGATCGGCCACCCCGCTCGCTCGCCATCTCGGCT 1380
DB 6437 GATTATGACAAACAGCATGAGCAGGATCGGCCACCCCGCTCGCTCGCCATCTCGGCT 6496
QY 1381 TTTTCTGTTCTCCGAGCAAAATGATGTAATTTGGCTGTCCCTCACTGACGCCAGTATGAC 1440

DB 6497 TTTTCTGTTCTCCGAGCAAAATGATGTAATTTGGCTGTCCCTCACTGACGCCAGTATGAC 6556
QY 1441 CAGTCCACTACGGTGTCTCAACTGGCCGGTTTATATCTCGGACAGGTGACTTTGGTG 1500
DB 6557 CAGTCCACTACGGTGTCTCAACTGGCCGGTTTATATCTCGGACAGGTGACTTTGGTG 6616
QY 1501 AATGTTGCGACTGGCGCGCAGGCCCTAGCCGATCGCTTACTGCTCAAAAGTCACCCCTC 1560
DB 6617 AATGTTGCGACTGGCGCGCAGGCCCTAGCCGATCGCTTACTGCTCAAAAGTCACCCCTC 6676
QY 1561 GACGGCGCGCCCTCCGACTGTTGAGCAATATCCAAAGATCTTGTGTGCTCCCCCTT 1620
DB 6677 GACGGCGCGCCCTCCGACTGTTGAGCAATATCCAAAGATCTTGTGTGCTCCCCCTT 6736
QY 1621 CGTGCAAGCTCTCCCTTTTGGGAGCGCGCAACAACAAAGAGGTTATCTTATAATAT 1680
DB 6737 CGTGCAAGCTCTCCCTTTTGGGAGCGCGCAACAACAAAGAGGTTATCTTATAATAT 6796
QY 1681 AATACTACTGCTAGTGACAGATCTGATTTGAAAATGCTGCCGGCATCGGGTCGCCATT 1740
DB 6797 AATACTACTGCTAGTGACAGATCTGATTTGAAAATGCTGCCGGCATCGGGTCGCCATT 6856
QY 1741 TCAACTATATACACAGGCTTGGGCGCGGTCCGCTCGCATTTCTGCGGCCCGGGTTTG 1800
DB 6857 TCAACTATATACACAGGCTTGGGCGCGGTCCGCTCGCATTTCTGCGGCCCGGGTTTG 6916
QY 1801 GCTCCACGCTCCGCCCTGCTCTGCTGGAGGATACCTTTTATTATCCGGGCGGGCGCAC 1860
DB 6917 GCTCCACGCTCCGCCCTGCTCTGCTGGAGGATACCTTTTATTATCCGGGCGGGCGCAC 6976
QY 1861 ACATTTGATGACTCTTCCGCTGAATGCGCGCTTTAGGCTCCAGGGTTGTGTTCCAG 1920
DB 6977 ACATTTGATGACTCTTCCGCTGAATGCGCGCTTTAGGCTCCAGGGTTGTGTTCCAG 7036
QY 1921 TCAACTGCTGCTAGCTCCAGCGCTTAAAGTTAAGTGGGTGAAACCTCGGAGTTGTAG 1980
DB 7037 TCAACTGCTGCTAGCTCCAGCGCTTAAAGTTAAGTGGGTGAAACCTCGGAGTTGTAG 7096

RESULT 12
US-09-128-275A-10
; Sequence 10, Application US/09128275A
; Patent No. 6229005
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbrough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fy, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,275A
; FILING DATE: 03-AUG-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:

QY	61	CGACCGGCTCAGCGCTCTGGCGCGCGTCTGGCGCGCGAGCGGCTACCGCGGCTGCT	120	Db	1141	GGCGGGCTCCCGACAGAAATTAATTTCTGTCGGCTGGCGGCAACTGTTTATTTCCGCGCG	1200
Db	61	CGACCGGCTCAGCGCTCTGGCGCGCGTCTGGCGCGCGAGCGGCTACCGCGGCTGCT	120	QY	1198	GTTCGTCTCAGCCAAATGCGGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAG	1257
QY	121	TTCTCGGGGTGACCGGGTGAATTTCTCAGCGCTTTCGCAATCCCTATATATTCATCAACCAAC	180	Db	1201	GTTCGTCTCAGCCAAATGCGGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAG	1260
Db	121	TTCTCGGGGTGACCGGGTGAATTTCTCAGCGCTTTCGCAATCCCTATATATTCATCAACCAAC	180	QY	1258	CAGGATAAGGGTGTTCCTATCCCCACGATATCGATCTTTGGTGATTCGCGTGTGGTCAAT	1317
QY	181	CCCTTTGCCCGAGAGAGTTGCCGCTGCTGCGGGTCTGGACCTCGCTCGCCAAACGAGCC	240	Db	1261	CAGGATAAGGGTGTTCCTATCCCCACGATATCGATCTTTGGTGATTCGCGTGTGGTCAAT	1320
Db	181	CCCTTTGCCCGAGAGAGTTGCCGCTGCTGCGGGTCTGGACCTCGCTCGCCAAACGAGCC	240	QY	1318	CAGGATAATGACAAACGAGCATGAGGATTCGGGCCAACCCCGTGGCTGGCGCAATCTCGG	1377
QY	241	CGGCGACTTGGCTCCACTTTGGCGAGATCAGGCGCCGCTCGCTCGCTCGCTCGCTCGC	300	Db	1321	CAGGATAATGACAAACGAGCATGAGGATTCGGGCCAACCCCGTGGCTGGCGCAATCTCGG	1380
Db	241	CGGCGACTTGGCTCCACTTTGGCGAGATCAGGCGCCGCTCGCTCGCTCGCTCGCTCGC	300	QY	1378	CTTTTCTGTTCTCCGAGCAATGATGTAATTTGGCTGTCCCTACTGAGCCGAGTAT	1437
QY	301	CGACCTGCCACAGCGCGGGCTCGCGGCTGACGGCTGTGGCGCTGSCCATGACACCTCA	360	Db	1381	CTTTTCTGTTCTCCGAGCAATGATGTAATTTGGCTGTCCCTACTGAGCCGAGTAT	1440
Db	301	CGACCTGCCACAGCGCGGGCTCGCGGCTGACGGCTGTGGCGCTGSCCATGACACCTCA	360	QY	1438	GACCACTCCACTTACGGGTCTCACTGGCCCGGTTTATATCTCGGACAGCGTGAATTTG	1497
QY	361	CCGCTCCCGGAGCTTGAATTTCTCGGGTGCATTTCTACGCGCGCAGTATATTTGTCTACT	420	Db	1441	GACCACTCCACTTACGGGTCTCACTGGCCCGGTTTATATCTCGGACAGCGTGAATTTG	1500
Db	361	CCGCTCCCGGAGCTTGAATTTCTCGGGTGCATTTCTACGCGCGCAGTATATTTGTCTACT	420	QY	1498	GTGAATTTGCGACTGCGCGCGAGCGCTAGCCGATCGCTTGAAGTGTCCAAAGTCAAC	1557
QY	421	TCACCCCTGACATCTCTGTGGCTCTGCGCTCAATTTAGTCTGTATGCGAGCCCGCTT	480	Db	1501	GTGAATTTGCGACTGCGCGCGAGCGCTAGCCGATCGCTTGAAGTGTCCAAAGTCAAC	1560
Db	421	TCACCCCTGACATCTCTGTGGCTCTGCGCTCAATTTAGTCTGTATGCGAGCCCGCTT	480	QY	1558	CTGAGCGGGCGGCGCTCCCGACTGTTGAGCAATATTTCAAGACATTTCTTGTGTCTCCCG	1617
QY	481	AATCGGCTCTCGCGCTGCGGAGCGGTACTAATCTACATATATGCGCAGAGGCTCC	540	Db	1561	CTGAGCGGGCGGCGCTCCCGACTGTTGAGCAATATTTCAAGACATTTCTTGTGTCTCCCG	1620
Db	481	AATCGGCTCTCGCGCTGCGGAGCGGTACTAATCTACATATATGCGCAGAGGCTCC	540	QY	1618	CTTCTGTCGCAAGCTCTCTTTTGGAGGCGCGCACAAACAAAGCAGGTTATCCTTATAAT	1677
QY	541	AATATGACAGTACCGGGTTCGCGGCTACTATCGCTTACCGGCGCTTGTGCTTAAT	600	Db	1621	CTTCTGTCGCAAGCTCTCTTTTGGAGGCGCGCACAAACAAAGCAGGTTATCCTTATAAT	1680
Db	541	AATATGACAGTACCGGGTTCGCGGCTACTATCGCTTACCGGCGCTTGTGCTTAAT	600	QY	1678	TATAATCTACTGCTAGTGACAGATTTCTGATTGAAAATGCTGCGCGGCAATCGGGTCGCC	1737
QY	601	GCAGTTGAGGCTATGCTATATCCATTTCTTCTGGCTCAAAACCAACCAACCCCTACA	660	Db	1681	TATAATCTACTGCTAGTGACAGATTTCTGATTGAAAATGCTGCGCGGCAATCGGGTCGCC	1740
Db	601	GCAGTTGAGGCTATGCTATATCCATTTCTTCTGGCTCAAAACCAACCAACCCCTACA	660	QY	1738	ATTTCAACCTATACCAACAGGCTTGGGCGCGCTTGGCGGCTTCTGCGGCGCGGCTT	1797
QY	661	TTGTTGACATGAATTTCCATTTCTCCACTGATGTGAGGATTTCTTCACTGATTTGCGATA	720	Db	1741	ATTTCAACCTATACCAACAGGCTTGGGCGCGCTTGGCGGCTTCTGCGGCGCGGCTT	1800
Db	661	TTGTTGACATGAATTTCCATTTCTCCACTGATGTGAGGATTTCTTCACTGATTTGCGATA	720	QY	1798	TTGGCTCCACGCTCCGCGCTCTGCTGGAGGATATCTTTGATTTATCGGGCGGCGG	1857
QY	721	GCATCTGAATTTGCTATCCCAAGCGGCTTCACTACCGCAATCAAGGTTGCGGCTCG	780	Db	1801	TTGGCTCCACGCTCCGCGCTCTGCTGGAGGATATCTTTGATTTATCGGGCGGCGG	1860
Db	721	GCATCTGAATTTGCTATCCCAAGCGGCTTCACTACCGCAATCAAGGTTGCGGCTCG	780	QY	1858	CACACATTTGATGACTTCTGCGCTGAAATGCGCGCTTTAGGCTTCCAGGGTGTGCTTTC	1917
QY	781	GTGAGACATCTGCTGTGCTGAGGAGGAGCCACCTTCTGCTGCTGCTGCTGCTGCTGCTG	840	Db	1861	CACACATTTGATGACTTCTGCGCTGAAATGCGCGCTTTAGGCTTCCAGGGTGTGCTTTC	1920
Db	781	GTGAGACATCTGCTGTGCTGAGGAGGAGCCACCTTCTGCTGCTGCTGCTGCTGCTGCTG	840	QY	1918	CAGTCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAAGTGGGTAAAGCTCGGGAGTTG	1977
QY	841	CATGGCTCTCCAGTTAACTCTATACCAATACCCCTTATACCGGTCCTTGGCTTACTG	900	Db	1921	CAGTCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAAGTGGGTAAAGCTCGGGAGTTG	1980
Db	841	CATGGCTCTCCAGTTAACTCTATACCAATACCCCTTATACCGGTCCTTGGCTTACTG	900	QY	1978	TAG 1980	
QY	901	GACTTTGCTTACAGCTTGAATTTCCCAATCTCAGCAGCTTCAACCAATACAGCTGCTG	960	Db	1981	TAG 1983	
Db	901	GACTTTGCTTACAGCTTGAATTTCCCAATCTCAGCAGCTTCAACCAATACAGCTGCTG	960				
QY	961	TCCGCTTACTTCCAGCACTGCTGCTCAC- - -TCCGCGCGAGGGCGGAGCTGCGGAG	1017				
Db	961	TCCGCTTACTTCCAGCACTGCTGCTCAC- - -TCCGCGCGAGGGCGGAGCTGCGGAG	1020				
QY	1018	CTGACCAACATGCGAGCCACAGGTTCAAGAGATCTCCACTTACCGGCTTAAATGGG	1077				
Db	1021	CTGACCAACATGCGAGCCACAGGTTCAAGAGATCTCCACTTACCGGCTTAAATGGG	1080				
QY	1078	GTAGTGAAGTCCGCGGAGTACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA	1137				
Db	1081	GTAGTGAAGTCCGCGGAGTACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA	1140				
QY	1138	GGCGGGCTCCCGACAGAAATTAATTTCTGTCGGCTGGCGGCAACTGTTTATTTCCCGCGG	1197				

RESULT 15
PCT-US95-13703-2
; Sequence 2, Application PC/TUS9513703
; GENERAL INFORMATION:
; APPLICANT: HEPATITIS E VIRUS ANTIGENS AND
; TITLE OF INVENTION: USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA


```
QY 1678 TATATACTACTGCTAGTCACACAGATTCTGATTGAAATGCTGCCGCCATCGGTCGCC 1737
Db |||||
QY 1681 TATAATATCTACTGCTAGTCACACAGATTCTGATTGAAATGCTGCCGCCATCGGTCGCC 1740
Db |||||
QY 1738 ATTTCAACCTATACCAACAGGCTTGGGGCCGGTCCGGTCGCCATTTCTGCCGCCCGGGTT 1797
Db |||||
QY 1741 ATTTCAACCTATACCAACAGGCTTGGGGCCGGTCCGGTCGCCATTTCTGCCGCCCGGGTT 1800
Db |||||
QY 1798 TTGGCTCCACGCTCCGCCCTGGCTCTGCTGGAGGATACCTTTTGATTATCCGGGGCGGGCG 1857
Db |||||
QY 1801 TTGGCTCCACGCTCCGCCCTGGCTCTGCTGGAGGATACCTTTTGATTATCCGGGGCGGGCG 1860
Db |||||
QY 1858 CACACATTTGATGACTTCTGCCCTGAATGCCCGCTTTAGGCTCCAGGGTTGTGCTTTC 1917
Db |||||
QY 1861 CACACATTTGATGACTTCTGCCCTGAATGCCCGCTTTAGGCTCCAGGGTTGTGCTTTC 1920
Db |||||
QY 1918 CAGTCAACTGTCGCTGAGCTCCAGCGCCTTAAAGTTAAGTGGGTAAACTCGGGAGTTG 1977
Db |||||
QY 1921 CAGTCAACTGTCGCTGAGCTCCAGCGCCTTAAAGTTAAGTGGGTAAACTCGGGAGTTG 1980
Db |||||
QY 1978 TAG 1980
Db |||
QY 1981 TAG 1983
```

Search completed: August 21, 2004, 18:11:08
Job time : 167.394 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 15:15:13 ; Search time 885,329 Seconds
(without alignments)
10985.769 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096

Perfect score: 1980

Sequence: 1 ATGGCGCTAGGCTCTTTT.....GTAAACTCGGAGTGTAG 1980

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3228839 seqs, 245606551 residues

Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1980	100.0	2055	15	US-10-165-868-4
2	1980	100.0	2100	15	US-10-165-868-2
3	1980	100.0	7171	10	US-09-851-410-10
4	1980	100.0	7180	17	US-10-239-090A-49
5	1957.4	98.9	2058	9	US-09-769-066-2
6	1621.4	81.9	1647	9	US-09-769-066-4
7	1371.8	69.3	1990	13	US-10-381-770-5
8	1368.6	69.1	2049	9	US-09-769-066-1
9	1368.6	69.1	2049	15	US-10-165-868-3
10	1368.6	69.1	2094	15	US-10-165-868-1
11	1368.6	69.1	7195	10	US-09-851-410-6
12	1361.4	68.8	1989	13	US-10-381-770-6
13	1319.6	66.6	7256	17	US-10-239-090A-44
14	1315.2	66.4	7277	10	US-09-468-147-164

15	1315.2	66.4	7277	10	US-09-468-147-165	Sequence 165, Appl
16	1315.2	66.4	7277	13	US-10-319-745-164	Sequence 164, Appl
17	1315.2	66.4	7277	13	US-10-319-745-165	Sequence 165, Appl
18	1314	66.4	7228	17	US-10-239-090A-47	Sequence 47, Appl
19	1313.2	66.3	7233	17	US-10-239-090A-45	Sequence 45, Appl
20	1308	66.1	1308	15	US-10-165-868-10	Sequence 10, Appl
21	1301.2	65.7	7229	17	US-10-239-090A-46	Sequence 46, Appl
22	1301.2	65.7	7234	17	US-10-239-090A-11	Sequence 11, Appl
23	1278	64.5	7230	17	US-10-239-090A-48	Sequence 48, Appl
24	1274	64.3	7202	10	US-09-468-147-89	Sequence 89, Appl
25	1274	64.3	7202	10	US-09-468-147-90	Sequence 90, Appl
26	1274	64.3	7202	13	US-10-319-745-89	Sequence 89, Appl
27	1274	64.3	7202	13	US-10-319-745-90	Sequence 90, Appl
28	1090.2	55.1	1647	9	US-09-769-066-3	Sequence 3, Appl
29	984	49.7	984	9	US-09-769-066-6	Sequence 6, Appl
30	981	49.5	981	15	US-10-165-868-8	Sequence 8, Appl
31	880.6	44.5	1311	15	US-10-165-868-9	Sequence 9, Appl
32	673.6	34.0	984	9	US-09-769-066-5	Sequence 5, Appl
33	673.6	34.0	984	15	US-10-165-868-7	Sequence 7, Appl
34	660.8	33.4	1020	10	US-09-468-147-129	Sequence 129, Appl
35	660.8	33.4	1020	13	US-10-319-745-129	Sequence 129, Appl
36	640.6	32.4	1389	10	US-09-468-147-193	Sequence 193, Appl
37	640.6	32.4	1389	10	US-09-468-147-196	Sequence 196, Appl
38	640.6	32.4	1389	13	US-10-319-745-193	Sequence 193, Appl
39	640.6	32.4	1389	13	US-10-319-745-196	Sequence 196, Appl
40	638.2	32.2	1026	10	US-09-468-147-192	Sequence 192, Appl
41	638.2	32.2	1026	10	US-09-468-147-195	Sequence 195, Appl
42	638.2	32.2	1026	13	US-10-319-745-192	Sequence 192, Appl
43	638.2	32.2	1026	13	US-10-319-745-195	Sequence 195, Appl
44	558.2	28.2	899	10	US-09-468-147-24	Sequence 24, Appl
45	558.2	28.2	899	13	US-10-319-745-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-10-165-868-4
; Sequence 4, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; Two, Jr., Shin
; Purdy, Michael A.
; Tam, Albert W.
; Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10165.868
; FILING DATE: 06-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870.985A
; FILING DATE: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921

:	FILING DATE:	13-OCTOBER-1989
:	APPLICATION NUMBER:	US 367,486
:	FILING DATE:	16-JUNE-1989
:	APPLICATION NUMBER:	US 336,672
:	FILING DATE:	11-APRIL-1989
:	APPLICATION NUMBER:	US 208,997
:	FILING DATE:	17-JUNE-1988
:	ATTORNEY/AGENT INFORMATION:	
:	NAME:	Fabian, Gary R.
:	REGISTRATION NUMBER:	33,875
:	REFERENCE/DOCKET NUMBER:	4600-0093.30
:	TELECOMMUNICATION INFORMATION:	
:	TELEPHONE:	(415) 324-0880
:	TELEFAX:	(415) 324-0960
:	INFORMATION FOR SEQ ID NO: 4:	
:	SEQUENCE CHARACTERISTICS:	
:	LENGTH:	2055 base pairs
:	TYPE:	nucleic acid
:	STRANDEDNESS:	double
:	TOPOLOGY:	linear
:	MOLECULE TYPE:	CDNA TO mRNA
:	HYPOTHETICAL:	NO
:	ORIGINAL SOURCE:	
:	INDIVIDUAL ISOLATE:	ORF 2, MEXICO, FIGURE 7
:	SEQUENCE DESCRIPTION:	SEQ ID NO: 4:
:	US-10-165-868-4	
:	Query Match	100.0%; Score 1980; DB 15; Length 2055;
:	Best Local Similarity	100.0%; Pred. No. 0;
:	Matches 1980; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	ATGCGCCTAGGCCTCTTTTGTCTGTGTTCCTCTTGTTTCTGCTATGTGCCCGGCCA 60
Dd	1	ATGCGCCTAGGCCTCTTTTGTCTGTGTTCCTCTTGTTTCTGCTATGTGCCCGGCCA 60
Qy	61	CCGACCGGTCAAGCGTCTGGCGCGCTGTGGCGGCGCACGCGCGGTACCGCGGTGT 120
Dd	61	CCGACCGGTCAAGCGTCTGGCGCGCTGTGGCGGCGCACGCGCGGTACCGCGGTGT 120
Qy	121	TTCGGGGTGACCGGGTGAATCTCAGCCCTTCGCAATCCCTCATATTCATCCAACCAC 180
Dd	121	TTCGGGGTGACCGGGTGAATCTCAGCCCTTCGCAATCCCTCATATTCATCCAACCAC 180
Qy	181	CCCTTTGCCCCAGACGTTTGCCTGCGCTCGGCTCTGGACCTCGCCTTCGCCAACACGC 240
Dd	181	CCCTTTGCCCCAGACGTTTGCCTGCGCTCGGCTCTGGACCTCGCCTTCGCCAACACGC 240
Qy	241	CGBGCATTTGACTCCACTTGGGAGATCAGGCCCCAGCGCCCTCGCGTCCCTCGGTGCG 300
Dd	241	CGBGCATTTGACTCCACTTGGGAGATCAGGCCCCAGCGCCCTCGCGTCCCTCGGTGCG 300
Qy	301	CGACTGTCACACGCGGGCTGCGCGCTGACGGCTGTGGCGCTTGCCTGCCCATGACACCTCA 360
Dd	301	CGACTGTCACACGCGGGCTGCGCGCTGACGGCTGTGGCGCTTGCCTGCCCATGACACCTCA 360
Qy	361	CCCGTCCGAGCGTTGATTTCTCGGGTGAATTTCTACGCGCGCAGTAATAATTGTCTACT 420
Dd	361	CCCGTCCGAGCGTTGATTTCTCGGGTGAATTTCTACGCGCGCAGTAATAATTGTCTACT 420
Qy	421	TCACCCTTGACATCCTCTGTGGCCTTCGCACTAAATTTAGTCTCTGTATGACGCCCCCTT 480
Dd	421	TCACCCTTGACATCCTCTGTGGCCTTCGCACTAAATTTAGTCTCTGTATGACGCCCCCTT 480
Qy	481	AATCCGCTCTGCCGCTGCAGACGGTACTAATCTCACTATGACGAGCGCTCC 540
Dd	481	AATCCGCTCTGCCGCTGCAGACGGTACTAATCTCACTATGACGAGCGCTCC 540
Qy	541	AATTATGACAGTACCGGGTTCGCCGCTACTATCCGTTACCGCCCTAGTGCTAAT 600
Dd	541	AATTATGACAGTACCGGGTTCGCCGCTACTATCCGTTACCGCCCTAGTGCTAAT 600
Qy	601	GCA GTTGAGGCTATGCTATATCCATTTCTTCTGGCCTCAAACAACCAACCCCTACA 660

QY 901 GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACCACTGTGTAACCAATACACGTGTG 960
Db 946 GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACCACTGTGTAACCAATACACGTGTG 1005
QY 961 TCCCGTTACTTCAGCACTGCTGTCACCTCCGCGGAGGGCGCAGCGGAGCTGCGGAGCTG 1020
Db 1006 TCCCGTTACTTCAGCACTGCTGTCACCTCCGCGGAGGGCGCAGCGGAGCTGCGGAGCTG 1065
QY 1021 ACCACAATGCGAGCCACCAAGGTTGTAAGAAATCTCCATTTCACCGGCTTAATGGGTA 1080
Db 1066 ACCACAATGCGAGCCACCAAGGTTGTAAGAAATCTCCATTTCACCGGCTTAATGGGTA 1125
QY 1081 GGTGAAGTCGGCGGCGGATAGCTCTAACATTACTTAACCTTGTGACACAGCTCTCTCGAC 1140
Db 1126 GGTGAAGTCGGCGGCGGATAGCTCTAACATTACTTAACCTTGTGACACAGCTCTCTCGAC 1185
QY 1141 GGGCTCCCGACAGAATAATTTGCTCGGCTGGCGGGCAACTGTTTATTCGCCCGCGTT 1200
Db 1186 GGGCTCCCGACAGAATAATTTGCTCGGCTGGCGGGCAACTGTTTATTCGCCCGCGTT 1245
QY 1201 GTCTCAGCAATGGCGAGCCAAACGTTGAAGCTCTATACATCAGTGGAGAAATCTCAGCAG 1260
Db 1246 GTCTCAGCAATGGCGAGCCAAACGTTGAAGCTCTATACATCAGTGGAGAAATCTCAGCAG 1305
QY 1261 GATAGGGTGTGCTATATCCCGACAGATATCGATCTTGGTGATTTCGGGTGTGGTCAATCAG 1320
Db 1306 GATAGGGTGTGCTATATCCCGACAGATATCGATCTTGGTGATTTCGGGTGTGGTCAATCAG 1365
QY 1321 GATTATGACAACAGCATGAGCAGGATCGGCCACCCCGTCCGCTCGCCCATCTCGGCT 1380
Db 1366 GATTATGACAACAGCATGAGCAGGATCGGCCACCCCGTCCGCTCGCCCATCTCGGCT 1425
QY 1381 TTTTCTGTTCTCGAGCAAAATGATGTAATTTGGTCTGCTCCTCACTCAGCGGATGATGAC 1440
Db 1426 TTTTCTGTTCTCGAGCAAAATGATGTAATTTGGTCTGCTCCTCACTCAGCGGATGATGAC 1485
QY 1441 CAGTCCACTTACGGGTGCTCACTGCGCGGTTTATATCTCGGACAGCGTGAATTTGGTG 1500
Db 1486 CAGTCCACTTACGGGTGCTCACTGCGCGGTTTATATCTCGGACAGCGTGAATTTGGTG 1545
QY 1501 AATGTTGCACTGGCGCGAGGCGGTAGCCCGATCGCTGACTGTGTCGCAAGTCAACCTC 1560
Db 1546 AATGTTGCACTGGCGCGAGGCGGTAGCCCGATCGCTGACTGTGTCGCAAGTCAACCTC 1605
QY 1561 GACGGCGGCGCTCCGACTGTGAGCAATATTCAGACATTTCTTGTCCTCCGCTT 1620
Db 1606 GACGGCGGCGCTCCGACTGTGAGCAATATTCAGACATTTCTTGTCCTCCGCTT 1665
QY 1621 CGTGGCAAGCTCTCTTTTGGAGCGCGGACACAAAGCAGGTTATCTTATAATTAT 1680
Db 1666 CGTGGCAAGCTCTCTTTTGGAGCGCGGACACAAAGCAGGTTATCTTATAATTAT 1725
QY 1681 AATACTACTGCTAGTGACCAATTTGATTGAAATGCTCGCGGCCATCGGGTCGCCATT 1740
Db 1726 AATACTACTGCTAGTGACCAATTTGATTGAAATGCTCGCGGCCATCGGGTCGCCATT 1785
QY 1741 TCAACTATACACAGGCTTGGCGCGGTCCGCTCGCCATTTCTGGCGCGGCTTTG 1800
Db 1786 TCAACTATACACAGGCTTGGCGCGGTCCGCTCGCCATTTCTGGCGCGGCTTTG 1845
QY 1801 GCTCCACGCTCCGCGCTGCTGCTGGAGGATCTTTGATTATCCGGGCGGCGCAC 1860
Db 1846 GCTCCACGCTCCGCGCTGCTGCTGGAGGATCTTTGATTATCCGGGCGGCGCAC 1905
QY 1861 ACATTTGATGACTTTCGCTCGTAATCGCGGCTTTAGGCTTCAGGTTGTGCTTCCAG 1920
Db 1906 ACATTTGATGACTTTCGCTCGTAATCGCGGCTTTAGGCTTCAGGTTGTGCTTCCAG 1965
QY 1921 TCAACTCTGCTGAGCTCCAGCGCTTAAAGTTAGGTGGTAAACTCTCGGAGTTGTAG 1980
Db 1966 TCAACTCTGCTGAGCTCCAGCGCTTAAAGTTAGGTGGTAAACTCTCGGAGTTGTAG 2025

RESULT 3
US-09-851-410-10
; Sequence 10, Application US/09851410
; Publication No. US20030124510A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; Yarbough, Patrice O
; Bradley, Daniel W
; Krawczynski, Krzysztof Z
; Tam, Albert
; Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; No. US20030124510A1-A/No. US20030124510A1-B Hepatitis Viral
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,410
; FILING DATE: 07-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/128,275
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0980
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1717 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: Composite Mexico strain
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-410-10

Query Match 100.0%; Score 1980; DB 10; Length 7171;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCGCGCCCTAGGCGCTCTTTGCTGTGTTCTCTTGTCTGCTATGTTGCCGCGCCA 60
Db 5117 ATCGCGCCCTAGGCGCTCTTTGCTGTGTTCTCTTGTCTGCTATGTTGCCGCGCCA 5176
QY 61 CCGACCGGTGAGCGCTCTGCGCGCGCTGTTGGGCGCGCAGCGCGGTACCGCGGTGGT 120

Db 5177 CCGACCGGTACCGGTCTGGCGCGCTGTGGCGGCGACGGCGGTACCGCGGTGT 5236
QY 121 TTCTGGGGTGACCGGGTGTATCTTCAGCCCTTCGGAATCCCGCTATATTCATCCAAACCAAC 180
Db 5237 TTCTGGGGTGACCGGGTGTATCTTCAGCCCTTCGGAATCCCGCTATATTCATCCAAACCAAC 5296
QY 181 CCCTTTGCCCGCCAGAGCTTGGCGCTCGCTCGGGCTCGGACCTCGCTTCGCGCAACCAAGCC 240
Db 5297 CCCTTTGCCCGCCAGAGCTTGGCGCTCGCTCGGGCTCGGACCTCGCTTCGCGCAACCAAGCC 5356
QY 241 CGGCCACTTGGCTCCACTTGGCGGAGATCAGCGCCAGCGCCCTCGCTTCGCTCCCGTGGC 300
Db 5357 CGGCCACTTGGCTCCACTTGGCGGAGATCAGCGCCAGCGCCCTCGCTTCGCTCCCGTGGC 5416
QY 301 CGACTGCCACAGCGGGGCTCGCGCGCTGAGCGGTGAGCGCTGCGGCTCGCCCATGACACTCA 360
Db 5417 CGACTGCCACAGCGGGGCTCGCGCGCTGAGCGGTGAGCGGTGCGGCTCGCCCATGACACTCA 5476
QY 361 CCCGTCGGGACGTTGATTTCTCGCGGTGCAATTTCTAGCGGCCAGTATAATTTTGTCTACT 420
Db 5477 CCCGTCGGGACGTTGATTTCTCGCGGTGCAATTTCTAGCGGCCAGTATAATTTTGTCTACT 5536
QY 421 TCACCCCTGACATCCTCTGTGGCCCTCTGGCACTAATTTAGTCTGTATGCAAGCCCGCCT 480
Db 5537 TCACCCCTGACATCCTCTGTGGCCCTCTGGCACTAATTTAGTCTGTATGCAAGCCCGCCT 5596
QY 481 AATCCGCTCTCGCGCTGCAGGACGGTACTAATTAATCACTATATGCGCCACAGAGCGCTCC 540
Db 5597 AATCCGCTCTCGCGCTGCAGGACGGTACTAATTAATCACTATATGCGCCACAGAGCGCTCC 5656
QY 541 AATTAATGACAGTACCGGGTTGCGCGCTACTATCGTTTACCGGCCCTAGTGCCTAAT 600
Db 5657 AATTAATGACAGTACCGGGTTGCGCGCTACTATCGTTTACCGGCCCTAGTGCCTAAT 5716
QY 601 GCAAGTTGAGGCTAATGCTATATCAATTTCTTTCTGGCCCTCAAAACCAACCAACCCCTACA 660
Db 5717 GCAAGTTGAGGCTAATGCTATATCAATTTCTTTCTGGCCCTCAAAACCAACCAACCCCTACA 5776
QY 661 TCTGTTGACATGAATTCATTAATCTTCACTCATGTCAAGATCTTGTTCACACCTGGGATA 720
Db 5777 TCTGTTGACATGAATTCATTAATCTTCACTCATGTCAAGATCTTGTTCACACCTGGGATA 5836
QY 721 GCATCTGAATTTGGTCAATCCCAAGCGAGCGCTTCACTACCGCAATCAAGGTTGGCGTCC 780
Db 5837 GCATCTGAATTTGGTCAATCCCAAGCGAGCGCTTCACTACCGCAATCAAGGTTGGCGTCC 5896
QY 781 GTTGAGACATCTGGTGTGCTGAGAGGAAAGCCACCTCCGGTCTTGTCAATGTTATGCATA 840
Db 5897 GTTGAGACATCTGGTGTGCTGAGAGGAAAGCCACCTCCGGTCTTGTCAATGTTATGCATA 5956
QY 841 CATGGCTCTCCAGTTAATCTTACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 900
Db 5957 CATGGCTCTCCAGTTAATCTTACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 6016
QY 901 GACTTTTGCCTTAGAGCTTAGTTTTCGCAATCTCACCACCTGTAAACCAACCAACCAACGTTG 960
Db 6017 GACTTTTGCCTTAGAGCTTAGTTTTCGCAATCTCACCACCTGTAAACCAACCAACCAACGTTG 6076
QY 961 TCCCGTTACTCCAGCACTGCTCGTCACTCCGCGCGAGGGGCGGAGGACTCGGAGCTG 1020
Db 6077 TCCCGTTACTCCAGCACTGCTCGTCACTCCGCGCGAGGGGCGGAGGACTCGGAGCTG 6136
QY 1021 ACCAACAATGACGACAGGTTTCAATGAAGATCTCACTTTACCGGCTTAATGGGGTA 1080
Db 6137 ACCAACAATGACGACAGGTTTCAATGAAGATCTCACTTTACCGGCTTAATGGGGTA 6196
QY 1081 GGTGAAGTCCGCGCGGAGTCTTAACATTAATTAACCTTGTGTACAGCTCTCGGC 1140
Db 6197 GGTGAAGTCCGCGCGGAGTCTTAACATTAATTAACCTTGTGTACAGCTCTCGGC 6256
QY 1141 GGGCTCCCGACAGAAATTAATTTCTCGGCTGGCGGCAACTGTTTATTTCCCGCCCGGTT 1200

Db 6257 GGGCTCCCGACAGAAATTAATTTCTGTGCGCTGGCGGGCAACTGTTTATTTCCCGCCCGGTT 6316
QY 1201 GTCTCAGCCAAATGCGGAGCCAAACCGTGAAGCTCTATATCATCAGTGGAGAATGCTCAGCAG 1260
Db 6317 GTCTCAGCCAAATGCGGAGCCAAACCGTGAAGCTCTATATCATCAGTGGAGAATGCTCAGCAG 6376
QY 1261 GATAAGGGTGTGTCTATCCCCCAGATATCGATCTTGCTGATTCGCTGTCGTCATTCAG 1320
Db 6377 GATAAGGGTGTGTCTATCCCCCAGATATCGATCTTGCTGATTCGCTGTCGTCATTCAG 6436
QY 1321 GATTATGACAAACCAAGCATGAGCAGGATCGGCCCAACCCCGTCCGCTCGGCCACTCTCGGCT 1380
Db 6437 GATTATGACAAACCAAGCATGAGCAGGATCGGCCCAACCCCGTCCGCTCGGCCACTCTCGGCT 6496
QY 1381 TTTTCTGTTCTCCGAGCAATGATGATGTTGGCTGTCCCTCACTGAGCCGAGTATGAC 1440
Db 6497 TTTTCTGTTCTCCGAGCAATGATGATGTTGGCTGTCCCTCACTGAGCCGAGTATGAC 6556
QY 1441 CAGTCCACTTACGGGTCTCAACTGGCCCGGTTTATATCTCGGACAGCTGACTTTGGTG 1500
Db 6557 CAGTCCACTTACGGGTCTCAACTGGCCCGGTTTATATCTCGGACAGCTGACTTTGGTG 6616
QY 1501 AATGTTCCGACTGCGCGCAGCGCGCTAGCCGATCGCTTGACTGCTCCTCAAGTCAACCTC 1560
Db 6617 AATGTTCCGACTGCGCGCAGCGCGCTAGCCGATCGCTTGACTGCTCCTCAAGTCAACCTC 6676
QY 1561 GACGGGCGGCCCTCCCGACTGTTGAGCAATATTCGAAGACATTTCTTTGTGCTCCCGCT 1620
Db 6677 GACGGGCGGCCCTCCCGACTGTTGAGCAATATTCGAAGACATTTCTTTGTGCTCCCGCT 6736
QY 1621 CGTGGCAAGCTCTCTTTTGGGAGCGCGCACAAAGCAGGTTATCCTTATAATAT 1680
Db 6737 CGTGGCAAGCTCTCTTTTGGGAGCGCGCACAAAGCAGGTTATCCTTATAATAT 6796
QY 1681 AATACTACTGCTAGTGACCAATTTCTGAATGAAATGCTGCGGCCCATCGGGTCCGCAT 1740
Db 6797 AATACTACTGCTAGTGACCAATTTCTGAATGAAATGCTGCGGCCCATCGGGTCCGCAT 6856
QY 1741 TCAACCTATACCAAGCTTGGGCGCGGTTCGGTTCGCATTTCTGCGCGCGCGGTTTG 1800
Db 6857 TCAACCTATACCAAGCTTGGGCGCGGTTCGGTTCGCATTTCTGCGCGCGCGGTTTG 6916
QY 1801 GCTCCAGCTCCCGCTCGGCTCTGCTGAGGATCTTTGATTAATCCGGGCGGCGCAC 1860
Db 6917 GCTCCAGCTCCCGCTCGGCTCTGCTGAGGATCTTTGATTAATCCGGGCGGCGCAC 6976
QY 1861 ACATTTGATGACTTCTGCCCTGAATGCGCGCTTTAGGCCCTCCAGGTTGTGTTCCAG 1920
Db 6977 ACATTTGATGACTTCTGCCCTGAATGCGCGCTTTAGGCCCTCCAGGTTGTGTTCCAG 7036
QY 1921 TCAACTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACCTCGGAGTTGTAG 1980
Db 7037 TCAACTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACCTCGGAGTTGTAG 7096

RESULT 4

US-10-239-090A-49
; Sequence 49, Application US/10239090A
; Publication No. US20040101820A1
; GENERAL INFORMATION:
; APPLICANT: KABUSHIKI KAISHA TOSHIBA
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus col
; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same, i
; FILE REFERENCE: 02S0741P
; CURRENT APPLICATION NUMBER: US/10/239,090A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 57
; SEQ ID NO 49
; LENGTH: 7180
; TYPE: DNA

Db 1441 GACGAGTCCACTTACGGGTGTCACCTGGCCGGGTTTATATCTCGGACAGCGTGACTTTG 1500
Qy 1498 GTGAATGTGGACTGGGCGGAGCCGTAGCCGATCGCTTGATGCTGCTCCAAAGTCACC 1557
Db 1501 GTGAATGTGGACTGGGCGGAGCCGTAGCCGATCGCTTGATGCTGCTCCAAAGTCACC 1560
Qy 1558 CTGACGGGCGCCCTCCGACTCTGAGCAATATCCAGACATTTCTTTGTGCTCCCC 1617
Db 1561 CTGACGGGCGCCCTCCGACTCTGAGCAATATCCAGACATTTCTTTGTGCTCCCC 1620
Qy 1618 CTTCGTGGCAAGCTCTCTTTTGGAGGCGGCAACAACAAAGCAGGTTATCTTATAT 1677
Db 1621 CTTCGTGGCAAGCTCTCTTTTGGAGGCGGCAACAACAAAGCAGGTTATCTTATAT 1680
Qy 1678 TATAATCTACTGCTAGTACAGATTTCTGATTGAATGCTGCGGCGCATCGGTCGCC 1737
Db 1681 TATAATCTACTGCTAGTACAGATTTCTGATTGAATGCTGCGGCGCATCGGTCGCC 1740
Qy 1738 ATTTCAACTATACCAAGGCTTGGGCGGCTCGGTCGCCATTTCTCGGCGCGGTT 1797
Db 1741 ATTTCAACTATACCAAGGCTTGGGCGGCTCGGTCGCCATTTCTCGGCGCGGTT 1800
Qy 1798 TTGGCTCCAGCTCCGCGCTCTGCTGAGGATACCTTTTGAATATCCGGGCGGCG 1857
Db 1801 TTGGCTCCAGCTCCGCGCTCTGCTGAGGATACCTTTTGAATATCCGGGCGGCG 1860
Qy 1858 CACACATTTGAGCTTCTGCGCTGATGCGGCTTTAGGCGCTCCAGGTTGTGCTTTC 1917
Db 1861 CACACATTTGAGCTTCTGCGCTGATGCGGCTTTAGGCGCTCCAGGTTGTGCTTTC 1920
Qy 1918 CAGTCAACTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGTAAAGTCTCGGAGTTG 1977
Db 1921 CAGTCAACTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGTAAAGTCTCGGAGTTG 1980
Qy 1978 TAG 1980
Db 1981 TAG 1983

RESULT 6
US-09-769-066-4
; Sequence 4, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; McAttee, C. Patrick
; Yarbrough, Patrice O.
; Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/769,066
; FILING DATE: 24-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/542,634
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1647 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: Hepatitis E virus (Mexico strain)
; r62kDa, FIGURE 2
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-769-066-4

Query Match 81.9%; Score 1621.4; DB 9; Length 1647;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1638; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Qy 334 GCTGTGGGCGCTGCCCATGACACCTCACCCGTCGCCGAGCGTTGATTTCTCGCGTGCAATT 393
Db 1 GCTGTGGGCGCTGCCCATGACACCTCACCCGTCGCCGAGCGTTGATTTCTCGCGTGCAATT 60
Qy 394 CTACGCGCGCAGTATAATTTGTCTACTTCAACCCCTGACATCTCTGTGCGCTCTGCACT 453
Db 61 CTACGCGCGCAGTATAATTTGTCTACTTCAACCCCTGACATCTCTGTGCGCTCTGCACT 120
Qy 454 RAATTTAGTCTCTATGACGAGCGGCTTAAATCGCGCTCTGCGCTGCGAGCGGTACTAAT 513
Db 121 AATTTAGTCTCTATGACGAGCGGCTTAAATCGCGCTCTGCGCTGCGAGCGGTACTAAT 180
Qy 514 ACTCACATTATGCGCACAGAGCGCTCCAAATTAATGACAGTACCGGTTGCGCGGCTACT 573
Db 181 ACTCACATTATGCGCACAGAGCGCTCCAAATTAATGACAGTACCGGTTGCGCGGCTACT 240
Qy 574 ATCCGTTACCGGCGCTCTAGTCTTAATGAGTTGGAGGCTATGCTATATCAATTTCTTTC 633
Db 241 ATCCGTTACCGGCGCTCTAGTCTTAATGAGTTGGAGGCTATGCTATATCAATTTCTTTC 300
Qy 634 TGGCGCTCAACACACCAACCCCTACATCTGTTGACATGAATTCCTCAATTTCCACTGAT 693
Db 301 TGGCGCTCAACACACCAACCCCTACATCTGTTGACATGAATTCCTCAATTTCCACTGAT 360
Qy 694 GTCAGGATTTCTTTCAACCTGGCATAGCATCTGAATTTGGTATCCCAAGAGCGCGCTT 753
Db 361 GTCAGGATTTCTTTCAACCTGGCATAGCATCTGAATTTGGTATCCCAAGAGCGCGCTT 420
Qy 754 CACTACCGCAATCAAGTTGGCGCTCGGTTGAGACATCTGGTGTGCTGAGGAGGAGGCC 813
Db 421 CACTACCGCAATCAAGTTGGCGCTCGGTTGAGACATCTGGTGTGCTGAGGAGGAGGCC 480
Qy 814 ACCTCGGCTTCTGTCATGTTATGACATACATGCTCTCCAGTTAACTTAACTTAACTTAACT 873
Db 481 ACCTCGGCTTCTGTCATGTTATGACATACATGCTCTCCAGTTAACTTAACTTAACTTAACT 540
Qy 874 CTTTATACCGGTGCGCTTGGCTTTACTGAGCTTTTGCCTTAGAGCTTGGTTTTCGAATCTC 933
Db 541 CTTTATACCGGTGCGCTTGGCTTTACTGAGCTTTTGGCTTAGAGCTTGGTTTTCGAATCTC 600
Qy 934 ACCACCTGTAACACCAATACACGTTGTCCTGCTTACTCCAGCACTGCTCGTCACT---TCC 990
Db 601 ACCACCTGTAACACCAATACACGTTGTCCTGCTTACTCCAGCACTGCTCGTCACT---TCC 660
Qy 991 GCCCGAGGGGCGCAGCGGACTGCGGAGCTGACCAAACTGACGACCAAGCTTCAAGAA 1050
Db 661 CGCCGAGGGGCGCAGCGGACTGCGGAGCTGACCAAACTGACGACCAAGCTTCAAGAA 720
Qy 1051 GATCTCCACTTTTACCGGCTTAAATGGGGTAGGTGAAGTCCGCGCGGGAATGCTCTAACA 1110
Db 721 GATCTCCACTTTTACCGGCTTAAATGGGGTAGGTGAAGTCCGCGCGGGAATGCTCTAACA 780
Qy 1111 TTACTTAACCTTGTGTGACAGCTCTCTCGGCGGCTCCCGACAGAAATTAATTTCTCGCT 1170

[illegible]

RESULT 7
US-10-381-770-5
; Sequence 5, Application US/10381770
; Publication No. US20040052813A1
; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPOS
; TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
; FILE REFERENCE: IEC010037PCT
; CURRENT APPLICATION NUMBER: US/10/381,770
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1

Db 241 CGACCACTCGGCTCGCTTGGCGTACACAGGCCCGCCCGCGCTTGCCTCAAGTCGT 300
Qy 301 CGACCTGCCACAGCCGGGCTCGCGCGCTGACGGCTGTGGCGCCCTGCCCATGACACCTCA 360
Db 301 AGACCTACCAAGCTGGGGCCGGCGGTAAACGGGTGCTCCGGCCCATGACACCCGG 360
Qy 361 CCGCTCCGGAGGTTGATTTCTCGCGGTGCAATTTCTACGGCCGCGAGTAAATTTGCTACT 420
Db 361 CCAGTGCCTGATGTGCACTCCGCGCGGCGCACTTTGGCGCGGCGAGTATAAACCCTATCA 420
Qy 421 TCACCCCTGACATCTCTGTGGCCCTCTGGCACTAATTTAGTCTGTATGACGCCGCCCTT 480
Db 421 TCTCCCTTACCTCTCTCGGTGGCCACCGGCATAACCTGGTTCTTTATGGCCGCCCTCTT 480
Qy 481 AATCCGCTCTGCCCTGACGAGCGGTACTAATCTACATTAATGCGCACAGAGCCCTCC 540
Db 481 AGTCGCTTTTACCCCTTCAGAGCGGCACCAATACCATAAATGAGCCAGGAGCTTCT 540
Qy 541 AATTATGACAGTACCGGTTGCCCGGCTACTATCCGTTACCGGCCCTTAGTGCTTAAT 600
Db 541 AATTATGCCCACTACCGGTTGCCCGGTGCCCAATCCGTTACCGGCCGTGGTCCCAAT 600
Qy 601 GCAGTTGAGGCTAGCTATATCCATTTCTTCTGGCCCTCAAAACACCAACCCCTACA 660
Db 601 GCTGTGGCGGTTAGCCCATCTCCATCTCATTTCTGGCCACAGACCACCAACCCCGAG 660
Qy 661 TCTGTGACATGAATTCATTAATTTCCACTGATGTGAGGATTTCTTTTCAACCTGGCATA 720
Db 661 TCCGTTGATATGAATTCATTAATTTCCAGCGATGTTCTGATTTAGTCCAGCCCGCATA 720
Qy 721 GCATCTGAATTTGTCATCCCAAGGAGCGCCTTCACTACCGCAATCAAGGTTGGCGCTCG 780
Db 721 GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGCTCC 780
Qy 781 GTTCAGACATCGGTGTGCTGAGGAGGAGCGACCTCCGGTCTTGTGATGTATGACATA 840
Db 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGCTGTGTTATGCTTTGCATA 840
Qy 841 CATGGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTACTG 900
Db 841 CATGGCTACTCGTAATTTCTATCTATCTATCTAATACCCCTATACCGGTGCCCTCGGGCTGTTG 900
Qy 901 GACTTTGCTTAGAGCTTGAATTTGCGAACTTACCCCGGTTAACCAATACCGGGCTC 960
Db 901 GACTTTGCGCTTAGAGCTTGAATTTGCGAACTTACCCCGGTTAACCAATACCGGGCTC 960
Qy 961 TCCGTTTACTCCAGACTGCTCGTCACTCC---GCCGAGGGCGGAGCGGACTGGGAG 1017
Db 961 TCCGTTTATTCAGACTGCTCGGACCGCCTTCTCGCGGTGGGAGCGGACTGGGAG 1020
Qy 1018 CTGACCAACTGACGACCGCAGGTTTCATGAAGATCTCCACTTTTACCGGCTTAATGGG 1077
Db 1021 CTCACCAACCGGCTGTACCCGCTTTATGAAGAGCTCTATTTACTAGTACTAATGGT 1080
Qy 1078 GTAGTGAAGTGGCGGGGATAGCTCTAATCTAATCTAATCTAATCTGTCGACGCTCCTC 1137
Db 1081 GTGCGTGAATCGGCGGGGATAGCTCTAATCTAATCTAATCTAATCTGTCGACCTGCTT 1140
Qy 1138 GCGGGCTCCGACAGATTAATTTGCTCGCTGGCGGCACTGTTTATTTCCGCGCG 1197
Db 1141 GCGGGCTCCGACAGATTAATTTGCTCGCTGGCGGCACTGTTTATTTCCGCGCTCC 1200
Qy 1198 GTTGTCTCAGCAATGGCGACCAACCGTGAAGCTCTATACATCAGTGGAGAACTCAG 1257
Db 1201 GTTGTCTCAGCAATGGCGACCACTGTTAAGTTGTATATCTGTAGAGAACTGCTCAG 1260
Qy 1258 CAGGATAAGGTTGTCTATCCCGACGATATGATCTTGGTGAATCGCGTGGTGCATT 1317
Db 1261 CAGGATAAGGTTATGCAATCCCGCATGACATTGACCTCGGAGAACTCTCGTGGTTATT 1320
Qy 1318 CAGGATTTGACACAGCATGACGAGATCGGCCACCCCGTGGCTGGCCACTCGG 1377
Db 1321 CAGGATTTATGAACCAACATGAACAAGATCGGCCGAGCGCTTCTCCAGCCCATCGCG 1380

Qy 1378 CTTTTTCTGTCTCGAGCAATGATGATCTTTGGCTGTCCCTCACTGCGCGGAGTAT 1437
Db 1381 CTTTTTCTGTCTCGAGCAATGATGATGCTTTGGCTGTCTCTCACTGCGCGGAGTAT 1440
Qy 1438 GACCACTGCTCACTTACGGGTGCTCAACTGCGCGGTTTATCTCGGACAGCGTACTTGG 1497
Db 1441 GACCACTGCTCACTTACGGGTGCTCAACTGCGCGGTTTATCTCGGACAGCGTACTTGG 1500
Qy 1498 GTGAATGTTGGAATGCGGCGGAGCCGCTAGCCGATCGCTTGAAGTCTGAAAGTCAAC 1557
Db 1501 GTTAATGTTGGAATGCGGCGGAGCCGCTAGCCGATCGCTTGAAGTCTGAAAGTCAAC 1560
Qy 1558 CTCGAGGGGCGGCCCTCCGACTGTTGAGCAATATTCAGAGCAATCTTTGTGTGCTCC 1617
Db 1561 CTGACGGTGGCGGCCCTCTCCACCATTCAGCAGTACTCGAAGACCTTCTTTGTCTGCG 1620
Qy 1618 CTTCGTGCAAGCTCTCCCTTTTGGAGGCGCGCACAAACAAAGCAGGTTATCTTTATAAT 1677
Db 1621 CTCGCGGTAGCTCTCTTTCTGGAGGCGAGCACAACTAAAGCCGGGTACCTTTATAAT 1680
Qy 1678 TATTAATCTACTGCTAGTGAACAGATCTGATGAAATGCTGCGGCGCATCGGTCGCC 1737
Db 1681 TATAACACCACTGCTAGCGCAACTGCTTCTCGAGAAATCGCGCGCACCGGTCGCT 1740
Qy 1738 ATTTCAACCTATACACAGGCTTGGGCGCGTCCGTCGCGCATTTCTGCGCGCGCGTT 1797
Db 1741 ATTTCCACTTACACCACTAGCTGCGTGGTCTGCTCCGCTCTCCATTTCTGCGGTTGCC 1800
Qy 1798 TTGGCTCCAGCTCGCGCTCGCTCTGCTGAGGATACTTTTGTGATTTATTCGCGGGCGGG 1857
Db 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTGCCCGCGCC 1860
Qy 1858 CACACATTTGATGATCTTCTGCCCTGAATGCGCGCTTTAGGCTCCAGGTTTGTCTTTC 1917
Db 1861 CATACTTTGATGATTTCTGCCAGAGTGCGCCCTTGCCCTTCAGGCTGCGCTTC 1920
Qy 1918 CAGTCACTGCTGCTGAGCTCCAGGCTTAAAGTTAAGTGGGTAAACCTCGGAGTTG 1977
Db 1921 CAGTCACTGCTGCTGAGCTTCAAGGCTTAAAGTTAAGTGGGTAAACCTCGGAGTTG 1980
Qy 1978 TAG 1980
Db 1981 TAG 1983

RESULT 9

US-10-165-868-3
; Sequence 3, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:

APPLICANT: Reyes, Gregory R.
Bradley, Daniel W.
Twu, Jr-Shin
Purdy, Michael A.
Tam, Albert W.
Krawczynski, Krzysztof Z.
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method

NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/165,868


```
Db 1561 CTTGACGCTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTCTTTGTCTGCG 1620
QY 1618 CTTGTGCAAGCTCTCTCTTTTGGAGCGCGCACAAACAAAGAGGTTATCTTATAAT 1677
Db 1621 CTCGCGGTAAAGCTCTCTTTCTGGAGGAGGACAACTAAAGCGGTACCTTTATAAT 1680
QY 1678 TATAATACTACTGTAGTGACACAGATTCTGATTGAAATGCTCGCGGCATCGGTGCGC 1737
Db 1681 TATAACACCACTGTAGGACCACTGCTTGTGAGAAATGCGCGGACCGGTGCTCT 1740
QY 1738 ATTTCAACCTATACACAGAGTTGGGCGCGTCCGTCGCGCATTTCTGCGCGCGGTT 1797
Db 1741 ATTTCCACTTACACACTAGCTGGGTGCTGCTCCGTCCTCATTTCTGCGGTTGCGGTT 1800
QY 1798 TTGCTCCAGCTCGCCCTGCTCTGCTGAGGATACCTTTGATTATCCGGGGGCGC 1857
Db 1801 TTAGCCCCCACTCGCGTAGCATTTGCTTGAGGATACCTTTGGACTACCTTGCCGCGCC 1860
QY 1858 CACACATTTGATGACTTCTGCGCTGAATGCGCGCTTTAGGCCTCCAGGTTGTGCTTTC 1917
Db 1861 CATACTTTGATGATTTCTGCCAGAGTGCGCGCCCTTGSCCTTACGGGCTGCGCTTTC 1920
QY 1918 CAGTCACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTTG 1977
Db 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAAGATGAAGTGGGTAAACTCGGAGTTG 1980
QY 1978 TAG 1980
Db 1981 TAG 1983
```

RESULT 10

```
US-10-165-868-1
; Sequence 1, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
;           Bradley, Daniel W.
;           Twu, Ji-Shin
;           Purdy, Michael A.
;           Tam, Albert W.
;           Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/165,868
; FILING DATE: 06-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; APPLICATION NUMBER: US 336,672
```

```
; FILING DATE: 11-APRIL-1989
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2094 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: BURMA SEQUENCE, FIGURE 7
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-165-868-1
```

Query Match 69.1%; Score 1368.6; DB 15; Length 2094;

Best Local Similarity 81.0%; Pred. No. 0; Matches 1606; Conservative 0; Mismatches 374; Indels 3; Gaps 1;

```
QY 1 ATGCGCCTAGGCCTCTTTTGTCTGTTGTTCTCTCTTGTTCCTCTATGTTCTGCTATGTTCCCGCGCCA 60
Db 46 ATGCGCCTCGGCCTATTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 105
QY 61 CCGACCCGCTAGCCGCTCTGCGCGCGCTGCTGGGGCGGCGAGCGCGGTGACCGGGGTGGT 120
Db 106 CCGCGCGCTCAGCCGCTCTGCGCGCGCTGCTGGGGCGGCGAGCGCGGTGCTGGGGGTGGT 165
QY 121 TCTGCGGCTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCCTCAACCAAC 180
Db 166 TCTGCGGCTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCCTCAACCAAC 225
QY 181 CCCTTTGCCAGACGTTGCGGCTGCTGGGGTCTGGAGCTCGCCCTTCGCCAACCAAGCC 240
Db 226 CCCTTTCGCCCGCATGTGTCAGCGCTGCGCGCGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 285
QY 241 CGGCACTTGGCTCCACTTGGCGAGATCAGGCCCGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 286 CGACCACTCGGCTCGCTTGGCGGTGACGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 345
QY 301 CGACCTGCCACAGCGCGGCTGCGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCT 360
Db 346 AGACCTACACAGCTGGGCGCGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCT 405
QY 361 CCGTCCCGGACGTTGATTCTCGCGGTGCAATTTACGCCGCCAGTATAATTTGCTACT 420
Db 406 CCAGTGCCTGATGTCGACTCCCGCGGCGCATCTTGGCGCGGAGTATAACCTATCAACA 465
QY 421 TCACCCCTGACATCCTCTGCGGCTCTGCGCAATAATTTAGTCTGATGAGCGCCCTT 480
Db 466 TCTCCCTTACCTCTTCCGTGGCCACCGGCATAAACCCTGCTTCTTTATGCGCGCTCTT 525
QY 481 AATCCGCTCTGCGGCTGCGAGCGGTACTTAATCTACATTCACATTAATGCGCACAGAGGCTCC 540
Db 526 AGTCCGCTTTTACCTCTTCCAGCGGACCAATACCATATTAATGCGCACAGAGGCTTCT 585
QY 541 AATTATGCACAGTACCGGGTTGCCGCTACTATCCGTTACCGCGCCCTAGTGCCTAAT 600
Db 586 AATTATGCCAGTACCGGGTTGCCGCTGCGCAATCCGTTACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
QY 601 GCAGTTGGAGGCTATGCTATATTCATTCTTTCTGGCTCAAAACCAACCAACCCCTACA 660
Db 646 GCTGTGCGCGGTTACCGCATCTCATCTATTCTGGCCACAGACCAACCAACCCCGAGC 705
QY 661 TCTGTTGACATGAATTCOAATTTACTTCCACTGATGTCAGGATTTCTTCAACCTGCTGATA 720
```

Db 706 TCCGTTGATGAAATCAATAAACCCTCGACGAGTGTTCGTATTTTAGTCCAGCCCGGCATA 765
Qy 721 GCATCTGAATTGGTCAATCCCAAGCAGCGCCTTCACTACCGCAATCAAGGTTGGGCTCG 780
Db 766 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCCTACACATATCGTAACCAAGGCTGGCGCTCC 825
Qy 781 GTTCGAGACATCTGGTGTGTGAGGAGGAAGCCACCTCCGCTCTGTGTCATGTTATGCATA 840
Db 826 GTCGAGACCTCTGGGTGGCTGAGGAGGAGGTACCTCTGCTCTGTGTTATGCTTTGCATA 885
Qy 841 CATGGCTCTCCAGTTAACTCCCTATACCAATACCCCTTATACCGGTCCTTGGCTTACTG 900
Db 886 CATGGCTCACTCGTAAATTCCTATATACTAATAACACCTATACCGGTGCTCGGCTGTTG 945
Qy 901 GACTTTGCTTAGAGCTTGAGTTTCGGAATCTCACCACCTGTAAACACCAATACAGCTGTG 960
Db 946 GACTTTGCTTAGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC 1005
Qy 961 TCCCGTTACTCAGCACTGCTGTCACCTCC--GCGCAGGGGCGCAGCGGACTCGCGAG 1017
Db 1006 TCCCGTTACTCAGCACTGCTGTCGCAACCGCCTTCGTCGGGTGCGGACGGGACTGCGAG 1065
Qy 1018 CTGACCACAACTGACGACCAACAGGTTCAAGAAATCTCCACTTACCGGCCCTTAATGGG 1077
Db 1066 CTCACCAACAGGCTGCTACCGCCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1125
Qy 1078 GTAGTGAAGTCGCGCGGGATAGCTCTAACTTACTTAACTTGTGACACGCTCCTC 1137
Db 1126 GTCGTGAGATCGCGCGGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTT 1185
Qy 1138 GCGCGGCTCCCGACAGAAATTAATTTGTCGCGTGGCGGCAACTGTTTATTTCCGCGCG 1197
Db 1186 GCGCGCTTCCGACAGAAATGATTTGTCGCGTGGTGGCCAGCTGTCTACTCCCGTCC 1245
Qy 1198 GTTGTCTCAGCAATGCGGAGCAACCGTGAAGCTCTATACATCAGTGGAGAACTGCTAG 1257
Db 1246 GTTGTCTCAGCAATGCGGAGCGCACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1305
Qy 1258 CAGGATAGGGTGTGCTATATCCCGACGATATCGATCTTGGTGATTCGGGTGTGGTCATT 1317
Db 1306 CAGGATAGGGTATTGCAATCCCGCATGACATTTGACCTCGAGAAATCTCGTGTGGTTATT 1365
Qy 1318 CAGGATATTGACAAACAGCATGACGAGATCGGCCACCCCGTCGCTCGCCATCTCGG 1377
Db 1366 CAGGATATTGATAACCAACATGAACAAGATCGCGGAGCCTTCTCCAGCCCATCGCGC 1425
Qy 1378 CTTTTTCTGTCTCGAGCAAAATGATGTACTTTGGCTGCTCTCTCTCACCGCTCGCGAGTAT 1437
Db 1426 CTTTTTCTGTCTCGAGCTAATGATGTGCTTTGGCTCTCTCTCTCACCGCTCGCGAGTAT 1485
Qy 1438 GACCACTCCACTACGGGTCGTCACACTGCGCGGTTTATATCTCGACAGCGTGACTTTG 1497
Db 1486 GACCACTCCACTATTGGCTCTTCGACTGGCCGACGTTTATGTTTCTGACTCTGTGACCTTG 1545
Qy 1498 GTGAATGTTGCGACTGCGCGCAGCGCGTAGCCCGATCGCTTGACTGTGCCAAAGTCACC 1557
Db 1546 GTTAATGTTGCGACCGCGCGCAGCGCGTTGCCGCTGCTCGATTGGACCAAGGTGACA 1605
Qy 1558 CTCGACGGGGGCGCTCCGACTGTGTAGCAATATTCAAGACATTTCTTGTGCTCCCG 1617
Db 1606 CTTGACGGTGGCGGCTCTCCCACTCCAGCAGTACTCGAAGACCTTCTTTGCTCGCGC 1665
Qy 1618 CTTGTTGCGCAAGCTCTCTTTTGGGAGCGCGGCACAAACAAAGCAGGTTATCTTATAAT 1677
Db 1666 CTCGCGGTTAGCTCTCTTTCTGGGAGGAGGACCAACTAAGCCGGTACCTTTATAAT 1725
Qy 1678 TATAATTACTGCTAGTGAACAGATTCTGATTGAAAAATGCTGCGCGGCATCGGGTCGCC 1737
Db 1726 TATAACACCACCTGCTAGCGACCAACTGCTGTGAGAAATGCCCGCGGACACCGGGTCGCT 1785
Qy 1738 ATTTCAACCTATACCAACAGCTTGGGCGCGGTCCGCTCGCCATTTCTCGGCGCGGGTT 1797
Db 1786 ATTTCCACTTACACCACTAGCCCTGGGGTGTGGTCCCGCTCTCCATTTCTCGCGGTTGCCGTT 1845

Qy 1798 TTGGCTCCACGCTCGGCCCTGGCTCTGCTGGAGGATACTTTTGTGATTATCCCGGGGCGGGCG 1857
Db 1846 TTAGCCCCCACTCTCGCTAGCATTTGCTTGAGGATACCTTGGACTACCTCGCCGCGGCC 1905
Qy 1858 CACCAATTTGATGACTTTCGCCCTGAATGCGCGCTTTAGGCTTCCAGGTTGTGCTTTC 1917
Db 1906 CATATTTTGTATGATTTCTGCCAGAGTGGCGCCCTTGGCCTTCAGGGCTGCGCTTTC 1965
Qy 1918 CAGTCAACTGTCGCTGAGCTCCAGCGCTTAAAGTTAAGGTGGGTAAAACTCGGAGTTG 1977
Db 1966 CAGTCTACTGCTGAGCTTCAGCGCTTAAAGTGAAGGTGGGTAAAACTCGGAGTTG 2025
Qy 1978 TAG 1980
Db 2026 TAG 2028

RESULT 11
US-09-851-410-6
; Sequence 6, Application US/09851410
; Publication No. US20030124510A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; Yardbough, Patrice O
; Bradley, Daniel W
; Krawczynski, Krzysztof Z
; Tam, Albert
; Fy, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,410
; FILING DATE: 07-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/128,275
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7195 base pairs
; TYPE: nucleic acid

Db	1621	CTCGCGGTAAGCTCTCTTTTGGAGACGAGTACTACTAAAGCCGGGTACCTCTTATAAT	1680
QY	1678	TATAATACTACTGCTAGTAGACAGATTCTGATTGAAAATGCTGCGGCCATCGGGTCGCC	1737
Db	1681	TATAACACCACTGCTAGTGACCACTGCTGTTGAGATCCGCTGGGCATCGGGTTGCT	1740
QY	1738	ATTTCACCTATACACAGAGCTTGGGCGCGGTCCGGTCGCATTTCCTGCGGCCGGGT	1797
Db	1741	ATTTCACCTATTACACCACTAGCTCGGGTCTGGTCCCGTCTCTATTTCGCGGTTCGTGT	1800
QY	1798	TTGGCTCCAGCTCCGCCCTGGCTCTGCTGGAGGATACCTTTTGATATCCGGGGCGGGG	1857
Db	1801	TTAGC-CCCCCTCCGCGCTAGCAATGCTTGAGGATACCATGGACTACCTCTCCCGCGC	1859
QY	1858	CACACATTTGATGACTTCTGCGCTCGAATGCCCGCTTTTAGGCTCCAGGCTTGTGCTTTC	1917
Db	1860	CATACCTTCGATGACTTCTGCCCGAGTGC CGCCCTTGGCTTCAGGGCTGTGCTTTT	1919
QY	1918	CAGTCACTGCTGAGTCCAGCGCTTAAAGTTAAAGTGGGTGAAACTCGGGAGTTG	1977
Db	1920	CAGTCTACTGCTGAGCTTCAGCGCTTAAAGTGAAGTGGGTAAAACTCGGGAGTTA	1979
QY	1978	TAG 1980	
Db	1980	TAG 1982	

RESULT 13
 US-10-239-090A-44
 ; Sequence 44, Application US/10239090A
 ; Publication No. US20040101820A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KABUSHIKI KAISHA TOSHIBA
 ; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co
 ; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same,
 ; TITLE OF INVENTION: detecting hepatitis E virus using the same
 ; FILE REFERENCE: 02S0741P
 ; CURRENT APPLICATION NUMBER: US/10/239,090A
 ; CURRENT FILING DATE: 2003-07-24
 ; PRIOR APPLICATION NUMBER: JP2001-191837
 ; PRIOR FILING DATE: 2001-06-25
 ; NUMBER OF SEQ ID NOS: 57
 ; SEQ ID NO 44
 ; LENGTH: 7256
 ; TYPE: DNA
 ; ORGANISM: Hepatitis E Virus Japan JKN-Sap (AB074918)
 US-10-239-090A-44

Query Match	66.6%;	Score 1319.6;	DB 17;	Length 7256;
Best Local Similarity	79.5%;	Pred. No. 0;		
Matches 1575;	Conservative	0;	Mismatches 404;	Indels 3; Gaps 1;

QY	1	ATGCGCCCTAGGCCTCTTTTGTGCTGTTGTTCTCTGTTTCTGCTATGTTGCCCGGCCA	60
Db	5190	ATGCGCCCTAGGCCTGTTCTGTGTTGTTCTCTGCTGTTTGGCTATGCTGCCCGGCCA	5249
QY	61	CCGACCGGTACGCCCTCTGCGCGCGCTGTTGGGCGCGGCGCGGTACCGGGGTGGT	120
Db	5250	CCGCGCGGTACGCCCTCTGCGCGCGTCCCGTGGGCGGCGAGCGCGGTGCGGGGTGGT	5309
QY	121	TTCTGGGTGACCGGGTTGATTCTACGCCCTTCGCAATCCCTATATTTCATCCAAACAAC	180
Db	5310	TTCTGGGTGACAGGGTTGATTCTACGCCCTTCGCCCTCCCTATATTTCATCCAAACAAC	5369
QY	181	CCCTTTGCCACAGAGTTGCGCGGTGCGGTCTGGACTCGCCTTCGCCCAACCGGCC	240
Db	5370	CCCTTCGCGCGCGATGTGCTTTCAAAACCGGGGCTGGAATCGCCCTCGACAGCCGCC	5429
QY	241	CGGCCACTTGGCTCCACTTGGGAGATCAGGCCAGGCCCTCCGCTGCTCCGTCGC	300
Db	5430	CGCCCCCTTGGCTCCGCTTGGGTGACCAAGTCCAGGCCCTCCGCTGCCCCCGCGT	5489
QY	301	CGACTGTCACAGCCGGGGCTCGGGCGCTGACGGCTGTGGCGCTGCCCATGACACTA	360

Db	5497	CGATCTGCCCGAGCTGGGGCTGGCGCTGACTGCCGTGTCAACCGCTCCTGACACAGCC	5556
Qy	361	CCCGTCCCGGACGTTGATTTCTCGGGTGCAATTTCTACGCCGCCAGATATAATTTGTCTACT	420
Db	5557	CCTGTACTGATTTGACTACAGTGGTCTATTCTCGCGCGGACGATACAAATTTGTCCACG	5616
Qy	421	TCACCCCTGACATCTCTGTGGCTCTGGCACTAAATTTAGTCCCTGATGACGCCCCCTT	480
Db	5617	TCCCGCTCAGCTCATCTGTGCTTCGGGTACTAAATTTGGTCCCTATGTGCCCCCGTG	5676
Qy	481	AATCGCCTCTGCGCTCGAGACCGGTACTAATACTCACATTTATGGCCACAGAGCCCTCC	540
Db	5677	AATCCCCCTCTCTCCAGATGGTACCACTCATATTTATGCTACTGAGGCATCC	5736
Qy	541	AATATGACAGTACCGGTTGCCCGCTACTATTCGGTTACCGGCCCTAGTGCCTAAT	600
Db	5737	AATATGCCAGTATCGGTTGTTGAGCTACAATCCGTTATCGGCCGTGTGCGCAAT	5796
Qy	601	CGAGTTGAGGTATGCTATATCCATTTCTTCTGGCCTCAAAACACCAACCCCTACA	660
Db	5797	GCCGTTGGTGGTATGCCAATTTCCATTTCTTTCTGGCCCCAACTACACTACCCCTACT	5856
Qy	661	TCGTGTGACATCAATTCATTTACTTCCACTGATGTCAGGATTTCTTTTCAACTGGCATA	720
Db	5857	TCGTGATATGAATTTCTATTACTTCCACVGAITTAGGATTTGGTTCAGCCCGGTATT	5916
Qy	721	GCATCTGAATGGTATCCCAAGCGAGCCCTTCACTACCGCAATCAAGTTGGCGCTCG	780
Db	5917	GCTTCGAGCTAGTCAATCCCAAGTAGCGCTTCAATACCGTAATCAAGGCTGGCGCTCT	5976
Qy	781	GTGAGACATCTGGTGTCTGAGGAGGAGCCACCTCGGTCTTCTCATGTTATGCATA	840
Db	5977	GTGAGACACCGGTTGGTCTGAGGAGGAGGCTACTTCGGTCTGTAATGCTTGTGATT	6036
Qy	841	CATGGCTCTCCAGTTAACTCCTATACCAATFACCCCTTATACCGTGCCCTTGGCTTACTG	900
Db	6037	CATGGCTCTCCTGTTAACTTCTTACACTAATACACCTTACACTGGTGGCGCTTCTT	6096
Qy	901	GACTTTGCCCTTAGAGCTTGAGTTTCGAAATCTCACACCTGTAAACCAATACAGTGTG	960
Db	6097	GATTTTGCACTAGAGCTTGAATTTAGGAATTTGACACCCGGGAACCAACACCCCGTGT	6156
Qy	961	TCCCGTACTCCAGCACTGCTCCTCAC---TCCGCCCGAGGGCCGACCGGACTCGGGAG	1017
Db	6157	TCCCGGTATACCAAGCAGCCCGCCCGCGCTCGCGCTGGTGTCTGATGGGACTGTGAG	6216
Qy	1018	CTGACCACAACTGCAGCCACAGGTTTCATGAAGAATCTCCACTTTACCGGCCCTTAATGGG	1077
Db	6217	CTTTACTACACAGCAGCCACACGTTTCATGAAGGACCTGCACITTCGCTGGCAGGAATGGC	6276
Qy	1078	GTAGGTGAAGTCGGCCCGGGATAGCTTAACATTTACTTAACCTTGTGTGACAGCTCCTC	1137
Db	6277	GTTGGTGAGGTGGTCTGGTATCGCCCTGACACTGTTTCAATCTCGCTGATACGCTTCTC	6336
Qy	1138	GGCGGGCTCCGACAGAAATTAATTTCTGGCTGGCGGGCAACTGTTTTATTTCCCGCCCG	1197
Db	6337	GGCGGTTTACGACAGAAATTTGATTTCTGGCTGGGGGCCAACTGTTTTACTCCCGCCCG	6396
Qy	1198	GTTGTCTCAGCCAAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAG	1257
Db	6397	GTTGTCTCAGCCAAATGGCGAGCCAAACAGTAAAGTTATATACATCTGTTGAGAATGCGAG	6456
Qy	1258	CAGGATTAAGGTGTGTCTATCCCCCAGATATCGATCTTGGTGAATTCGGTGGTTCATT	1317
Db	6457	CAGACAAAGGGCATCACCAATTTCCATGATATAGACTGGGTGACTCCCGTGGGTATC	6516
Qy	1318	CAGGATTAATGACCAACCAATGAGCAGATCGGCCCAACCGCTCGCTGGGCCATCTCGG	1377
Db	6517	CAGGATTAATGATAACAGCATGAGCAAGACCGACTACTCCGTCACTGCCCTCTCGC	6576
Qy	1378	CCTTTTCTGTCTCCGAGCAAAATGATGATTTTGGTGTCCCTCACTGACGCCGATAT	1437
Db	6577	CCCTTCTCAGTTCTTCGTGGCAATGATGTTTGTGGCTTTCCCTCACTGCCGCTGATAT	6636
Qy	1438	GACCAGTCCACTTACGGGTCTCAACTGGCCCGGTTTATATCTCGGACAGCGTGACTTTG	1497
Db	6637	GACCAGACTACGATGGGTCTCCACCAACCTTATGTATGTCTCTGACACAGTTACGCTT	6696
Qy	1498	GTGAATGTTGGACTGGCGCAGCCGATAGCCGATCGCTGACTGTGCTCAAGGTCACC	1557
Db	6697	GTTAATGTTGGCTACTGTGTCTCAGGCTTTGGCCGCTCCCTTGTAAAGTTACT	6756
Qy	1558	CTCGACGGGGCGCCCTCCCGACTGTTGAGCAATTTCCAGACATTTCTTTGTGTCTCCC	1617
Db	6757	CTGGACGGCGCCGCTTACTACCAATTCAGCAGTATTTCTAGACATTTTAIGTTCTCCCG	6816
Qy	1618	CTTCTGTGCAAGCTCTCTCTTTTGGGAGGCCGACCAACAAAAGCAGGTTATCTCTTATAT	1677
Db	6817	CTCCGCGGAGGCTGCTCTTTTGGGAGGCTGGCAGACTAAGGCCGGCTTACCCTTACRAAT	6876
Qy	1678	TATAATATCTACTGTAGTGACAGATTTCTGATTGAATGCTGCCGCCATCGGTCGCC	1737
Db	6877	TATAATATCTACCGCTAGTGACCAAAATTTTGATTGAGAAATGCGCGCCGCCACCGTGTGCT	6936
Qy	1738	ATTTCAACCTATACCAACAGGCTTGGGGCCGCTCGGTCGCCCATTTCTGCGGCCGCGTT	1797
Db	6937	ATTTCCACTATACCACTAGCTTAGTGCCGCTCTACCTCGATCTCTGCGGTGCGCGTA	6996
Qy	1798	TTGGCTCCACGCTCCGCTCGCTCTGCTGAGGATATACTTTTGATTAATTCGCGGGCGGCG	1857
Db	6997	CTGGCTCCACACTCTGCCCTTCTTGAGGATCTATTTGATTTACCCCGCCGTCGCC	7056
Qy	1858	CACACTTTGATGACTTCTGCCCTGAATGCGCGCTTTAGGCTCCAGGGTTGTGCTTTC	1917
Db	7057	CATACTTTTGTGATTTTTCGCCGAGTCCGCTACCTAGGTTTGAGGGTTGTGCAATTC	7116
Qy	1918	CAGTCAACTGTCGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAAACTCGGGAGTTG	1977
Db	7117	CAGTCTACTATTGCTGAGCTCCAGCGTTTAAAAATGAAGTAGGTAAACCCCGGAGTCT	7176
Qy	1978	TA 1979	
Db	7177	TA 7178	

Search completed: August 21, 2004, 22:42:05
Job time : 893.329 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 13:02:51 ; Search time 57 Seconds
(without alignments)
3271.603 Million cell updates/sec

Title: US-09-851-410a-8

Perfect score: 3408

Sequence: 1 MRPRPILLLLMLPMLPAP.....QSTVAELQLKMKVGKTRRL 660

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3408	100.0	660	2	Aar38785 HEV ORF2
2	3408	100.0	660	2	Aar39306 Burma str
3	3408	100.0	660	2	Aar96089 Hepatitis
4	3408	100.0	660	2	Aaw35826 Hepatitis
5	3408	100.0	660	2	Aaw80197 Protein e
6	3408	100.0	660	2	Aaw93389 Human HEV
7	3408	100.0	660	3	Aab24120 Hepatitis
8	3408	100.0	660	4	Aab62523 HEV-Burma
9	3408	100.0	660	5	Aaol15699 Hepatitis
10	3408	100.0	660	7	AD024376 Hepatitis
11	3401	99.8	660	2	Aaw76369 Hepatitis
12	3401	99.8	660	2	Aaw71210 Protein e
13	3399	99.7	660	2	Aar70323 Hepatitis
14	3391	99.5	660	2	Aar91814 Hepatitis
15	3391	99.5	660	2	Aaw93386 Human HEV
16	3391	99.5	660	5	AB881670 Hepatitis
17	3391	99.5	660	6	ADA50060 Hepatitis
18	3391	99.5	660	7	ADB97807 Hepatitis
19	3386	99.4	660	5	ABJ04803 Hepatitis
20	3383	99.3	660	4	AAE12380 Hepatitis
21	3382	99.2	660	2	AAE51265 HEV strai
22	3382	99.2	660	2	AAW81520 Hepatitis
23	3382	99.2	660	2	AAW93395 Human HEV
24	3382	99.2	660	2	AAW93388 Human HEV
25	3377	99.1	660	5	ABJ04811 Hepatitis

ALIGNMENTS

RESULT 1

AAR38785
ID AAR38785 standard; protein; 660 AA.

XX AAR38785;

XX AC

DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 11-JAN-1994 (first entry)

XX HEV ORF2 protein.

DE Enterically transmitted non-A non-B hepatitis; ET-NANB; vaccine.

XX Hepatitis E virus; Burma strain.

XX Key Location/Qualifiers

FT Peptide 225..660

FT Peptide /label= C2

FT Peptide 334..660

FT Peptide /label= SG3

FT Peptide 613..660

FT Peptide /label= 406.3-2

XX WO9314116-A1.

XX 22-JUL-1993.

XX 15-JAN-1993; 93WO-US000459.

XX 17-JAN-1992; 92US-00822335.

XX 01-MAY-1992; 92US-00876941.

XX (GENE-) GENELABS TECHNOLOGIES INC.

XX (USSH) US SEC DEPT HEALTH.

XX Reyes GR, Bradley DW, Tam AW, Carl M;

XX WPI; 1993-243144/30.

XX N-PSDB; AAQ47129.

XX New immunogenic hepatitis E virus (HEV) peptide(s) - are from the ORF1,

XX ORF2 and ORF3 regions of HEV, useful as a vaccine against HEV infection.

XX Disclosure; Fig 8; 48pp; English.

XX Immunogenic hepatitis E virus (HEV) peptides are selected from the ORF1,

XX ORF2 and ORF3 regions of HEV. The peptides can be used in vaccines to

26 3374 99.0 660 2 AAW93390 Human HEV
27 3372 98.9 660 2 AAW93394 Human HEV
28 3370 98.9 660 2 AAW93391 Human HEV
29 3370 98.9 660 5 ABJ04822 Hepatitis
30 3370 98.9 660 5 ABJ04819 Hepatitis
31 3369 98.9 660 5 ABJ04807 Hepatitis
32 3367 98.8 660 2 AAW93392 Human HEV
33 3363 98.7 660 5 ABJ04820 Hepatitis
34 3361 98.6 660 2 AAR14619 Protein e
35 3360 98.6 660 5 ABJ04806 Hepatitis
36 3360 98.6 660 5 ABJ04809 Hepatitis
37 3353 98.4 660 5 ABJ04821 Hepatitis
38 3352 98.4 660 5 ABJ04805 Hepatitis
39 3351 98.3 660 5 ABJ04813 Hepatitis
40 3351 98.3 660 5 ABJ04816 Hepatitis
41 3350 98.3 660 5 ABJ04810 Hepatitis
42 3349 98.3 660 5 ABJ04818 Hepatitis
43 3336 97.9 660 2 AAW93393 Human HEV
44 3290 96.5 655 5 ABJ04804 Hepatitis
45 3221 94.5 660 2 AAR96090 Hepatitis

CC prevent infection by HEV. The antibodies can neutralise and block HEV
CC infection and can be used to prevent or treat HEV infection. The peptides
CC and antibodies can also be used as diagnostic reagents. (Updated on 25-
CC MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS
CC field)

XX	Sequence 660 AA;
XX	Query Match 100.0%; Score 3408; DB 2; Length 660;
XX	Best Local Similarity 100.0%; Pred. No. 5.3e-263;
XX	Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MRPRPILLLLMFLMPLPAPPQPSGRRRRSGSGGGFWGDRVDSQPPAIPYIHTN 60
Db	1 MRPRPILLLLMFLMPLPAPPQPSGRRRRSGSGGGFWGDRVDSQPPAIPYIHTN 60
QY	61 PFAPDVTAAAGAGRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPADHTP 120
Db	61 PFAPDVTAAAGAGRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPADHTP 120
QY	121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTTHIMATEAS 180
Db	121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTTHIMATEAS 180
QY	181 NYAQYRVARATIRYRPLVPNAVGGYAISISFPQTPTTPTSDVMNSITSDVRLVQPGI 240
Db	181 NYAQYRVARATIRYRPLVPNAVGGYAISISFPQTPTTPTSDVMNSITSDVRLVQPGI 240
QY	241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSGLVMLCIHGSVNSYTNTPYTALGGLL 300
Db	241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSGLVMLCIHGSVNSYTNTPYTALGGLL 300
QY	301 DFALFELFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db	301 DFALFELFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY	361 VGEIGRGIALTLFNLADTLGGLPTELISAGGQLFYSRPVVSANGETVKLYTSVENAQ 420
Db	361 VGEIGRGIALTLFNLADTLGGLPTELISAGGQLFYSRPVVSANGETVKLYTSVENAQ 420
QY	421 QDKGIALPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWLSLTAAY 480
Db	421 QDKGIALPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWLSLTAAY 480
QY	481 DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQOYSKTFVFLP 540
Db	481 DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQOYSKTFVFLP 540
QY	541 LRGLKLSFWEAGTTKAGYPYNNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db	541 LRGLKLSFWEAGTTKAGYPYNNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY	601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVKTREL 660
Db	601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVKTREL 660

RESULT 2
AAR39306
ID AAR39306 standard; protein; 660 AA.
XX AAR39306;
AC AAR39306;
XX
DT 25-MAR-2003 (revised)
DT 14-FEB-1994 (first entry)
XX
DE Burma strain HEV ORF2 putative virus capsid protein.
XX
XX Hepatitis E virus; vaccine; neutralising antibodies; infection; block;
XX
XX open reading frame; antibodies.
OS
OS Hepatitis E virus.
XX

PN	WO9314208-A2.
XX	22-JUL-1993.
PD	
XX	19-JAN-1993; 93WO-US000475.
PF	
XX	17-JAN-1992; 92US-00822335.
PR	
XX	20-APR-1992; 92US-00870985.
XX	
XX	(GENE-) GENELABS TECHNOLOGIES INC.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Reyes GR, Bradley DW, Twu J, Purdy MA, Tam AW, Krawczynski KZ;
XX	
XX	WFI; 1993-243223/30.
DR	N-PSDB; AAQ46813.
XX	
XX	Antigen and antibody vaccines against hepatitis E virus infection -
PT	contain peptide(s) derived from capsid protein C-terminal or antibodies
PT	against protein.
XX	
PS	Disclosure; Fig 7; 43pp; English.
XX	
CC	The sequence is that of the putative virus capsid protein encoded by
CC	Burma strain hepatitis E virus (HEV) open reading frame ORF2. This
CC	protein or peptide fragments of it may be used in a vaccine composition
CC	for immunising an individual against HEV. Antibodies raised against these
CC	peptides can also be used in such vaccines. (Updated on 25-MAR-2003 to
CC	correct PN field.)
XX	
XX	Sequence 660 AA;
QY	Query Match 100.0%; Score 3408; DB 2; Length 660;
Db	Best Local Similarity 100.0%; Pred. No. 5.3e-263;
QY	Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1 MRPRPILLLLMFLMPLPAPPQPSGRRRRSGSGGGFWGDRVDSQPPAIPYIHTN 60
QY	61 PFAPDVTAAAGAGRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPADHTP 120
Db	61 PFAPDVTAAAGAGRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPADHTP 120
QY	121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTTHIMATEAS 180
Db	121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTTHIMATEAS 180
QY	181 NYAQYRVARATIRYRPLVPNAVGGYAISISFPQTPTTPTSDVMNSITSDVRLVQPGI 240
Db	181 NYAQYRVARATIRYRPLVPNAVGGYAISISFPQTPTTPTSDVMNSITSDVRLVQPGI 240
QY	241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSGLVMLCIHGSVNSYTNTPYTALGGLL 300
Db	241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSGLVMLCIHGSVNSYTNTPYTALGGLL 300
QY	301 DFALFELFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db	301 DFALFELFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY	361 VGEIGRGIALTLFNLADTLGGLPTELISAGGQLFYSRPVVSANGETVKLYTSVENAQ 420
Db	361 VGEIGRGIALTLFNLADTLGGLPTELISAGGQLFYSRPVVSANGETVKLYTSVENAQ 420
QY	421 QDKGIALPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWLSLTAAY 480
Db	421 QDKGIALPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWLSLTAAY 480
QY	481 DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQOYSKTFVFLP 540
Db	481 DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQOYSKTFVFLP 540
QY	541 LRGLKLSFWEAGTTKAGYPYNNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db	541 LRGLKLSFWEAGTTKAGYPYNNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600

```

Db 541 LRGLSFWEAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAV 600
QY 601 LAPHSSALLLEDITLDYPARAHTFDPCPCPLGLOGCAFQSTVLAELORLKMVGKTREL 660
Db 601 LAPHSSALLLEDITLDYPARAHTFDPCPCPLGLOGCAFQSTVLAELORLKMVGKTREL 660

RESULT 3
AAR96089
ID AAR96089 standard; protein; 660 AA.
XX
AC AAR96089;
XX
DT 16-OCT-2003 (revised)
DT 06-AUG-1996 (first entry)
XX
DE Hepatitis E virus (Burma strain) capsid protein.
KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
KW capsid.
XX
OS Hepatitis E virus; Burma strain.
XX
PN WO9612807-A2.
XX
PD 02-MAY-1996.
XX
PF 23-OCT-1995; 95WO-US013703.
XX
PR 24-OCT-1994; 94US-00327952.
PR 13-OCT-1995; 95US-00542634.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;
XX
DR WPI; 1996-230608/23.
DR N-PSDB; AAT27107.
XX
PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
XX diagnostic reagents for determining HEV infection and in vaccines.
PS Disclosure; Page 78-80; 125pp; English.
XX
CC The putative capsid protein (AAR96089) of hepatitis E virus (HEV) Burma
CC strain is encoded by ORF-2 (AAT27107) of the virus. PCR amplification of
CC ORF-2 allows prodn. of capsid protein or of C-terminal regions of the
CC capsid protein (see also AAR96091, AAR96093 and AAR96095) and expression
CC in Spodoptera frugiperda Sf9 insect cells using a baculovirus vector
CC provides recombinant C-terminal regions (see also AAR96101 and AAR96103)
CC useful as diagnostic reagents and in vaccines. The HEV Mexico strain
CC capsid protein (AAR96090) may similarly be used. (Updated on 16-OCT-2003
XX to standardise OS field)
XX
SQ Sequence 660 AA;

Query Match
Best Local Similarity 100.0%; Score 3408; DB 2; Length 660;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRFILLLLMFLPMLPAPPQSGRRGRSGSGGFGWDRVDSQPPAIPYIHTPN 60
Db 1 MRPRFILLLLMFLPMLPAPPQSGRRGRSGSGGFGWDRVDSQPPAIPYIHTPN 60
QY 61 PFAPDVTAAAGAPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPTAVPAHDTP 120
Db 61 PFAPDVTAAAGAPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPTAVPAHDTP 120
QY 121 PVPDVSRCAILLRQYNLSTSLTSSVATGNTLVLYAAPLSPILLPQDGTNTHIMATEAS 180
Db 121 PVPDVSRCAILLRQYNLSTSLTSSVATGNTLVLYAAPLSPILLPQDGTNTHIMATEAS 180

```

```

QY 181 NYAQYRVARATIRYRPLVNAVGGYVAISISWPOTTTTPTSDMNSITSTDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGGYVAISISWPOTTTTPTSDMNSITSTDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNOQWRSVETSGVAEEATSLVMLCIHGSLVNSYNTPTYGALGLL 300
Db 241 ASELVIPSERLHYRNOQWRSVETSGVAEEATSLVMLCIHGSLVNSYNTPTYGALGLL 300
QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLAADTLGGLPTELISSAGGOLFYSRPVVSANGBPTVKLYTSVNAQ 420
Db 361 VGEIGRGIALTLFNLAADTLGGLPTELISSAGGOLFYSRPVVSANGBPTVKLYTSVNAQ 420
QY 421 QDKGIAIPHDIDILGESRVVIQDYNQHEQDRPTSPAPSPRPFVLRANDVLMLSLTAAY 480
Db 421 QDKGIAIPHDIDILGESRVVIQDYNQHEQDRPTSPAPSPRPFVLRANDVLMLSLTAAY 480
QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSIDMTKVLTDGRPLSTIQOYSKTFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSIDMTKVLTDGRPLSTIQOYSKTFVLP 540
QY 541 LRGLSFWEAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAV 600
Db 541 LRGLSFWEAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAV 600
QY 601 LAPHSSALLLEDITLDYPARAHTFDPCPCPLGLOGCAFQSTVLAELORLKMVGKTREL 660
Db 601 LAPHSSALLLEDITLDYPARAHTFDPCPCPLGLOGCAFQSTVLAELORLKMVGKTREL 660

```

RESULT 4

AAW35826

ID AAW35826 standard; protein; 660 AA.

XX

AC AAW35826;

XX

DT 17-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 26-FEB-1998 (first entry)

XX

DE Hepatitis E virus Burma strain protein from ORF2.

XX

KW Hepatitis E virus; Burma; Mexico; immunoassay; peptide antigen; antibody;

XX diagnosis; HEV.

OS Hepatitis E virus; - Burma strain.

XX

PN US5686239-A.

XX

PD 11-NOV-1997.

XX

PF 09-MAY-1994; 94US-00240049.

XX

PR 17-JUN-1988; 88US-00208997.

XX

PR 11-APR-1989; 89US-00336672.

XX

PR 16-JUN-1989; 89US-00367486.

XX

PR 13-OCT-1989; 89US-00420921.

XX

PR 05-APR-1990; 90US-00505888.

XX

PR 17-JAN-1992; 92US-00681078.

XX

PR 20-APR-1992; 92US-00822335.

XX

PR 01-MAY-1992; 92US-00876941.

XX

PA (GENE-) GENELABS TECHNOLOGIES INC.

XX

PI Tam AW, Reyes GR, Yarbough PO;

XX

DR WPI; 1997-558132/51.

DR N-PSDB; AAT96959.

XX	PT	Diagnosis of hepatitis E virus Burma and Mexico strain infection - by immunoassay with hepatitis E virus peptide antigens.
XX	PS	Disclosure; Col 39-42; 36pp; English.
XX	XX	A method has been developed for detecting hepatitis E virus (HEV) antibodies (Ab). The method comprises: (a) reacting a serum sample with a HEV peptide antigen; and (b) examining the peptide for the presence of bound Ab, where the presence of bound Ab indicates the presence of HEV Ab. The present sequence represents the protein from the open reading frame, ORF2, from HEV Burma strain. The method can be used to diagnose infection with the enterically transmitted non-A/non-B viral hepatitis agent HEV, specifically the HEV Burma and Mexico strains. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 17-OCT-2003 to standardise OS field)
XX	SQ	Sequence 660 AA;
XX	Query Match	100.0%; Score 3408; DB 2; Length 660;
XX	Best Local Similarity	100.0%; Pred. No. 5.3e-263;
XX	Matches 660;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MRPRPILLLLMFLPMLPAPPGPQSGRRRCRRSGGGGFWGDRVDSQPPAIPYIHPTN 60
DB	1	MRPRPILLLLMFLPMLPAPPGPQSGRRRCRRSGGGGFWGDRVDSQPPAIPYIHPTN 60
QY	61	PFAPDVTAAAGAGPRVROPAPLGSARDQAPAVASRRRPTTAGAAPTAVAPAHDT 120
DB	61	PFAPDVTAAAGAGPRVROPAPLGSARDQAPAVASRRRPTTAGAAPTAVAPAHDT 120
QY	121	PVPDVSRGAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
DB	121	PVPDVSRGAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
QY	181	NYAQYRVARATIRYRPLVPNAVGGVAISISFWPQTITTSVDMNSITSTVRIILVOPGI 240
DB	181	NYAQYRVARATIRYRPLVPNAVGGVAISISFWPQTITTSVDMNSITSTVRIILVOPGI 240
QY	241	ASELVIPISERLHYRNOGWRVETSGVAEEATSGLVMLCIHGSLVNSYNTPTTGALGL 300
DB	241	ASELVIPISERLHYRNOGWRVETSGVAEEATSGLVMLCIHGSLVNSYNTPTTGALGL 300
XX	Query Match	100.0%; Score 3408; DB 2; Length 660;
XX	Best Local Similarity	100.0%; Pred. No. 5.3e-263;
XX	Matches 660;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MRPRPILLLLMFLPMLPAPPGPQSGRRRCRRSGGGGFWGDRVDSQPPAIPYIHPTN 60
DB	1	MRPRPILLLLMFLPMLPAPPGPQSGRRRCRRSGGGGFWGDRVDSQPPAIPYIHPTN 60
QY	61	PFAPDVTAAAGAGPRVROPAPLGSARDQAPAVASRRRPTTAGAAPTAVAPAHDT 120
DB	61	PFAPDVTAAAGAGPRVROPAPLGSARDQAPAVASRRRPTTAGAAPTAVAPAHDT 120
QY	121	PVPDVSRGAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
DB	121	PVPDVSRGAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
QY	181	NYAQYRVARATIRYRPLVPNAVGGVAISISFWPQTITTSVDMNSITSTVRIILVOPGI 240
DB	181	NYAQYRVARATIRYRPLVPNAVGGVAISISFWPQTITTSVDMNSITSTVRIILVOPGI 240
QY	241	ASELVIPISERLHYRNOGWRVETSGVAEEATSGLVMLCIHGSLVNSYNTPTTGALGL 300
DB	241	ASELVIPISERLHYRNOGWRVETSGVAEEATSGLVMLCIHGSLVNSYNTPTTGALGL 300
QY	301	DFALELEPRNLTGNTNTRVRSYSTARHRLRGADGTAELTTTAATPFMKDLYFTSTNG 360
DB	301	DFALELEPRNLTGNTNTRVRSYSTARHRLRGADGTAELTTTAATPFMKDLYFTSTNG 360
QY	361	VGEIGRGIALTLFNLADTLGLPTELISSAGGOLFYSRPVVSANGEPYVLYTSVENAQ 420
DB	361	VGEIGRGIALTLFNLADTLGLPTELISSAGGOLFYSRPVVSANGEPYVLYTSVENAQ 420
QY	421	QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
DB	421	QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
QY	481	DQSTYGSSTGYPVYSDSVTLNVATGAQAVARSLDWTKVLTDGRPLSTIQYKSTFFVLP 540
DB	481	DQSTYGSSTGYPVYSDSVTLNVATGAQAVARSLDWTKVLTDGRPLSTIQYKSTFFVLP 540
QY	541	LRGKLSFWEAGTTKAGYPYNYNTTASDOLLVNAAGHRVAISTYTTSLGAGPVSISAVAV 600
DB	541	LRGKLSFWEAGTTKAGYPYNYNTTASDOLLVNAAGHRVAISTYTTSLGAGPVSISAVAV 600
QY	601	LAPHSALALLEDTLPARAHTFDDFCPECRPLGLQGCAGFQSTVAELQRLKMKVKTREL 660
DB	601	LAPHSALALLEDTLPARAHTFDDFCPECRPLGLQGCAGFQSTVAELQRLKMKVKTREL 660
XX	RESULT 5	
XX	AAW80197	
XX	ID	AAW80197 standard; protein; 660 AA.
XX	XX	
XX	AC	AAW80197;

QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTTAATRFMKDLYFTSTNG 360
DB 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNADTLGGPLPTELISSAGGOLFYSRPFVVSANGEPPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNADTLGGPLPTELISSAGGOLFYSRPFVVSANGEPPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFVSLRANDVILWLSLTAAY 480
DB 421 QDKGIAIPHDIIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFVSLRANDVILWLSLTAAY 480
QY 481 DQSTYGSSTGPPVVSVDSTLVNATGAQAVARSIDWTKVLDGRPLSTIQOYKTFPVLP 540
DB 481 DQSTYGSSTGPPVVSVDSTLVNATGAQAVARSIDWTKVLDGRPLSTIQOYKTFPVLP 540
QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
DB 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDTLDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGKTREL 660
DB 601 LAPHSALELLEDTLDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGKTREL 660
RESULT 6
AAW93389
ID AAW93389 standard; protein; 660 AA.
AC AAW93389;
XX
XX
DT 11-JUN-1999 (first entry)
XX
DE Human HEV ORF 2 protein from strain Burma.
XX
XX Swine hepatitis E virus; HEV; cross-reaction; antibody; human; therapy;
KW vaccine; immunise; infection; detection; diagnosis; prevention.
XX
XX Hepatitis E virus.
XX
XX WO9904029-A2.
PN
XX
XX 28-JAN-1999.
PF 17-JUL-1998; 98WO-US014665.
XX
PR 18-JUL-1997; 97US-0053069P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
PI Meng X, Emerson SU, Purcell RH;
XX
XX WPI; 1999-132270/11.
DR
XX
XX New isolated swine hepatitis E virus - used to develop products for the
PT diagnosis, prevention and treatment of hepatitis E virus infection in
PT mammals, particularly humans.
XX
PS Example 1; Fig 3A; 70pp; English.
XX
CC This invention describes a swine hepatitis E virus (HEV) and its natural
CC mutants which are capable of cross-reacting with antibodies reactive with
CC a human HEV strain or natural mutants. The HEV and the proteins can be
CC used in vaccines for immunising against HEV infection. The swine HEV can
CC be used in humans to prevent possible infection by human HEV. The swine
CC HEV can also be used as a therapeutic treatment for infection by other
CC strains of HEV. The swine HEV can also be used for the production of
CC antibodies which can be used in therapy, detection and diagnosis. The
CC products can also be used for determining the susceptibility of cells or
CC organs to infection with swine HEV. The swine HEV is particularly useful
CC for the development of agents for the prevention, treatment and detection
CC of human HEV because it is not a human virus and thus can be handled both

CC experimentally and clinically without fear of severe infection and/or
CC contamination
XX
SQ Sequence 660 AA;
Query Match 100.0%; Score 3408; DB 2; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.3e-263;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPRPILLLLMLPMLPAPPPQPSGRRRGRSGSGGGFWGDRVDSQFPAIPIYHPTN 60
DB 1 MRPRPILLLLMLPMLPAPPPQPSGRRRGRSGSGGGFWGDRVDSQFPAIPIYHPTN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQAPAVASRRRPTTAGAPLTAVAPADHTP 120
DB 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQAPAVASRRRPTTAGAPLTAVAPADHTP 120
QY 121 PVPDVSRGAILLRQYNLSTPLTSSVATGNTNLVLAAPLSPLLPDQDGNTHIMATEAS 180
DB 121 PVPDVSRGAILLRQYNLSTPLTSSVATGNTNLVLAAPLSPLLPDQDGNTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAISISFWPQTTTPTTSVDMNSITSDVRILVQPGI 240
DB 181 NYAQYVARATIRYRPLVNAVGGYAISISFWPQTTTPTTSVDMNSITSDVRILVQPGI 240
QY 241 ASELVIPSERLHYRNQWRSVETSGVAEBEATSGLMVLCIHGSLVNSYNTPTGTALGLL 300
DB 241 ASELVIPSERLHYRNQWRSVETSGVAEBEATSGLMVLCIHGSLVNSYNTPTGTALGLL 300
QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTTAATRFMKDLYFTSTNG 360
DB 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNADTLGGPLPTELISSAGOLFYSRPFVVSANGEPPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNADTLGGPLPTELISSAGOLFYSRPFVVSANGEPPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFVSLRANDVILWLSLTAAY 480
DB 421 QDKGIAIPHDIIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFVSLRANDVILWLSLTAAY 480
QY 481 DQSTYGSSTGPPVVSVDSTLVNATGAQAVARSIDWTKVLDGRPLSTIQOYKTFPVLP 540
DB 481 DQSTYGSSTGPPVVSVDSTLVNATGAQAVARSIDWTKVLDGRPLSTIQOYKTFPVLP 540
QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
DB 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDTLDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGKTREL 660
DB 601 LAPHSALELLEDTLDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGKTREL 660
RESULT 7
AAB24120
ID AAB24120 standard; protein; 660 AA.
XX
XX AAB24120;
AC
XX
DT 29-JAN-2001 (first entry)
XX
XX Hepatitis E virus Burma strain ORF2 protein sequence SEQ ID NO:8.
DE Hepatitis E virus; HEV; non-A/non-B viral hepatitis; viral protein;
KW immunological; diagnosis; hepatitis; infection; identification;
KW detection; immunoreactive; hepatotropic; antiinflammatory; virucide;
KW vaccine; antiviral; antigenic; antibody; antigen.
XX
OS Hepatitis E virus.
XX
XX US6120988-A.
PN
XX

PD 19-SEP-2000.
 XX
 PF 07-JUN-1995; 95US-00478507.
 XX
 PR 17-JUN-1988; 88US-00208997.
 PR 11-APR-1989; 89US-00336672.
 PR 16-JUN-1989; 89US-00367486.
 PR 13-OCT-1989; 89US-00420921.
 PR 05-APR-1990; 90US-00505888.
 PR 05-APR-1991; 91US-00681078.
 PR 25-JUL-1994; 94US-00279823.
 XX
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 PA (USGO) US GOVERNMENT.
 XX
 PI Yarbough PO, Krawczynski KZ, Fry KE, Bradley DW, Tam A, Reyes GR;
 DR N-PSDB; AAA99259.
 DR WPI; 2000-593712/56.
 DR N-PSDB; AAA99259.
 XX
 PR Identifying recombinant antigen immunoreactive with antibody induced by
 PT hepatitis E virus (HEV), for detecting HEV infection, comprises
 PT immunoreacting a polypeptide from an HEV genome with an HEV-positive
 PT antiserum.
 XX
 PS Claim 6; Col 63-68; 46pp; English.
 XX
 CC The present invention describes a method for identifying a recombinant
 CC antigen immunoreactive with a hepatitis E virus (HEV)-induced antibody.
 CC The method comprises producing a polypeptide derived from an HEV genome,
 CC immunoreacting the polypeptide with an HEV-positive antiserum and
 CC selecting the polypeptide as a recombinant antigen if the polypeptide
 CC reacts with the HEV-positive antiserum. The method is useful for
 CC identifying recombinant antigen immunoreactive with antibody induced by
 CC HEV. The enterically transmitted non-A/non-B hepatitis (ET-NANB) (also
 CC known as HDV)-specific fragments are useful for identifying ET-NANB-
 CC derived cDNAs, which contain additional sequence information, as primers
 CC for detecting ET-NANB viral genomic material in a patient sample, for the
 CC synthesis of polypeptides for use in immunoassays, and for identifying
 CC similar antigenic regions encoded by related viral strains, e.g. Burmese
 CC strain. The antigens are especially useful in the preparation of vaccine
 CC against ET-NANB infection. These antigens may further be used to prepare
 CC antibodies to ET-NANB virus particles for use directly as antiviral
 CC agents, and to produce antiserum designed for pre- or post-exposure
 CC prophylaxis. The present sequence represents a specifically claimed HEV
 CC Burma strain protein sequence for use in the present invention
 XX
 SQ Sequence 660 AA;
 Query Match 100.0%; Score 3408; DB 3; Length 660;
 Best Local Similarity 100.0%; Pred. No. 5.3e-263; Mismatches 0; Gaps 0;
 Matches 660; Conservative 0; Indels 0; Gaps 0;
 QY 1 MRPRPILLLLMLPMLPAPPQPSGRRRRGSGGGGFWGDRVDSQPPAIPYIHTN 60
 DB 1 MRPRPILLLLMLPMLPAPPQPSGRRRRGSGGGGFWGDRVDSQPPAIPYIHTN 60
 QY 61 PFAPDVTAAGAGRVQPARPLGSAWRDQORFAVASRRRPTTAGAPLTAVAPADHTP 120
 DB 61 PFAPDVTAAGAGRVQPARPLGSAWRDQORFAVASRRRPTTAGAPLTAVAPADHTP 120
 QY 121 PVPDVSRGATLRQYNLSPLTSSVATGNLVLYAAPLSPLPLQDGTTHIMATEAS 180
 DB 121 PVPDVSRGATLRQYNLSPLTSSVATGNLVLYAAPLSPLPLQDGTTHIMATEAS 180
 QY 181 NYAQVRVARATIRVPLVPNAVGGVAISISFWPQTPTTPTSDVMNSITSTVRIIVQGI 240
 DB 181 NYAQVRVARATIRVPLVPNAVGGVAISISFWPQTPTTPTSDVMNSITSTVRIIVQGI 240
 QY 241 ASELVIPSERLHYRNOGRWSVETSGVAEEATSGIWMLCIHGSLVNSVTNTPYTGALGLL 300
 DB 241 ASELVIPSERLHYRNOGRWSVETSGVAEEATSGIWMLCIHGSLVNSVTNTPYTGALGLL 300

QY 301 DFALELEFRNLTEGNTNTRVSRYSSTARHRLRRGADGTAELETTAAATFRMKDLYFTSTNG 360
 DB 301 DFALELEFRNLTEGNTNTRVSRYSSTARHRLRRGADGTAELETTAAATFRMKDLYFTSTNG 360
 QY 361 VGEIGRGIALTLFNADTLGGLPTELISSAGGQLFYSRPVVSANGEPPTVKLYTSVENAQ 420
 DB 361 VGEIGRGIALTLFNADTLGGLPTELISSAGGQLFYSRPVVSANGEPPTVKLYTSVENAQ 420
 QY 421 QDKGIALPHDIDIGESRVVIQDYDNOHQDRPTSPAPSRPFPVSLRANDVLMLSLTAAY 480
 DB 421 QDKGIALPHDIDIGESRVVIQDYDNOHQDRPTSPAPSRPFPVSLRANDVLMLSLTAAY 480
 QY 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVTLDGRPLSTIQOYSKTFFVLP 540
 DB 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVTLDGRPLSTIQOYSKTFFVLP 540
 QY 541 LRKLSFWEAGTTKAGYPYNYNTASDQLLVENAAHRVAISTYTTSLGAGPVSISAVAV 600
 DB 541 LRKLSFWEAGTTKAGYPYNYNTASDQLLVENAAHRVAISTYTTSLGAGPVSISAVAV 600
 QY 601 LAPHSALLALLEDTLDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660
 DB 601 LAPHSALLALLEDTLDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660
 RESULT 8
 AAB62523
 ID AAB62523 standard; protein; 660 AA.
 AC AAB62523;
 DT 23-JUL-2001 (first entry)
 DE HEV-Burma strain viral protein ORF2.
 KW HEV; enterically transmitted nonA/nonB viral hepatitis agent; viral;
 KM ET-NANB; infection; vaccine.
 OS Hepatitis E virus.
 XX US6229005-B1.
 XX 08-MAY-2001.
 PF 03-AUG-1998; 98US-00128275.
 PR 17-JUN-1988; 88US-00208997.
 PR 11-APR-1989; 89US-00336672.
 PR 16-JUN-1989; 89US-00367486.
 PR 13-OCT-1989; 89US-00420921.
 PR 05-APR-1990; 90US-00505888.
 PR 05-APR-1991; 91US-00681078.
 PR 25-JUL-1994; 94US-00279823.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 XX
 PI Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam AW;
 PI Fry KE;
 XX WPI; 2001-342705/36.
 DR N-PSDB; AAF83495.
 XX
 PT New DNA sequences of enterically transmitted non-A/non-B (ET-NANB)
 PT hepatitis viral agent, useful in diagnosing infection by an enterically
 PT transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.
 XX
 PS Disclosure; Col 63-68; 45pp; English.
 XX
 CC The invention relates to an isolated DNA comprising the genome of an
 CC enterically transmitted nonA/nonB (ET-NANB) viral hepatitis agent (also
 CC referred as HEV). The DNA sequences or their fragments are useful in
 CC preparing ET-NANB viral proteins and as probes for virus detection. These

CC are particularly useful in diagnosing infection by an enterically
CC transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.
CC The present sequence represents a ORF2 protein encoded by a ET-NANB viral
CC DNA sequence from HEV-Burma strain

XX
SQ Sequence 660 AA;
Query Match 100.0%; Score 3408; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.3e-263;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPAIPYIHPTN 60
Db 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPAIPYIHPTN 60
QY 61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQORPAVASRRRPTTAGAAPTAVAPAHDP 120
Db 61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQORPAVASRRRPTTAGAAPTAVAPAHDP 120
QY 121 PVPDVSRGAILLRQYLNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRGAILLRQYLNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVPNAVGGVAISISFPQTTTPTSDVMNSITSDVRILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVPNAVGGVAISISFPQTTTPTSDVMNSITSDVRILVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGALGLL 300
Db 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGALGLL 300
QY 301 DFALEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGTAIPHIDIGSRVVIQYDQHQEQDRTPSPAPSRPFSVLRLANDVLSLTAARY 480
Db 421 QDKGTAIPHIDIGSRVVIQYDQHQEQDRTPSPAPSRPFSVLRLANDVLSLTAARY 480
QY 481 DOSTYGSSTGPPYVSDSVTLNVNATGAQAVARSLDWTVTLDGRPLSTTIQQYSKTFEVL 540
Db 481 DOSTYGSSTGPPYVSDSVTLNVNATGAQAVARSLDWTVTLDGRPLSTTIQQYSKTFEVL 540
QY 541 LRKLSFWEAGTTKAGYPYNNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKLSFWEAGTTKAGYPYNNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELTDLDYPARAHTFDDPCPCPLGQCAFQSTVAELQRLKMKVKTREL 660
Db 601 LAPHSALELTDLDYPARAHTFDDPCPCPLGQCAFQSTVAELQRLKMKVKTREL 660

RESULT 9
AAOI5699
ID AAOI5699 standard; protein; 660 AA.
XX
AC AAOI5699;
DT
XT 08-NOV-2002 (first entry)
XX
DE Hepatitis E virus (Burma strain) ORF2 protein.
XX
KW HEV; enterically transmitted nonA/nonB hepatitis viral agent;
KW Burma strain; bile; ORF2.
XX
OS Hepatitis E virus.
XX
PN US6379891-B1.
XX

PD 30-APR-2002.
XX
PF 19-APR-2000; 2000US-00553427.
XX
PR 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-00336672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 15-APR-1991; 91US-00681078.
PR 25-JUL-1994; 94US-00279823.
PR 07-JUN-1995; 95US-00478507.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
XX Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam A, Fry KE;
PI N-PSDB; AAL50386.
XX
DR WPI; 2002-517277/55.
XX
XX Detecting the presence of enterically transmitted nonA/nonB hepatitis
PT viral (HEV) agents in bile samples from infected humans and monkeys using
PT polymerase chain reaction.
XX
PS Disclosure; Col 19-24; 61pp; English.
XX
CC The invention comprises a method for detecting the presence of
CC enterically transmitted nonA/nonB hepatitis viral (HEV) agents in a
CC sample and isolating HEV agents or nucleic acid fragments produced by the
CC agent. The method utilises PCR - using bile from a human or cynomolgus
CC monkey actively infected with HEV as a source of the agent. The method of
CC the invention is used for detecting the presence of a viral agent in a
CC sample of cultured cells infected with the agent and isolating
CC enterically transmitted nonA/nonB HEV agents or nucleic acid fragments
CC produced by the agent. The present amino acid sequence represents the
CC protein encoded by open reading frame 2 (ORF2) of a DNA sequence isolated
CC from a Burmese strain of the Hepatitis E virus
XX
SQ Sequence 660 AA;

Query Match 100.0%; Score 3408; DB 5; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.3e-263;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPAIPYIHPTN 60
Db 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPAIPYIHPTN 60
QY 61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQORPAVASRRRPTTAGAAPTAVAPAHDP 120
Db 61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQORPAVASRRRPTTAGAAPTAVAPAHDP 120
QY 121 PVPDVSRGAILLRQYLNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRGAILLRQYLNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVPNAVGGVAISISFPQTTTPTSDVMNSITSDVRILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVPNAVGGVAISISFPQTTTPTSDVMNSITSDVRILVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGALGLL 300
Db 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGALGLL 300
QY 301 DFALEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420

QY 421 QDKGIALPHDIDLGESRVVIQDYDNOHQEODRPTSPAPSRPFVLRANDVWLSLTAAEY 480
Db 421 QDKGIALPHDIDLGESRVVIQDYDNOHQEODRPTSPAPSRPFVLRANDVWLSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVLDGRPLSTIQOYSKTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVLDGRPLSTIQOYSKTFFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660
Db 601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660

RESULT 10
ADD24376
ID ADD24376 standard; protein; 560 AA.
XX AC ADD24376;
XX DT 15-JAN-2004 (first entry)
XX DE Hepatitis E virus strain Burma ORF 2.
XX KW non-A/non-B viral hepatitis; ET-NANB; HEV; vaccine.
XX OS Hepatitis E virus; strain Burma.
XX PN US2003124510-A1.
XX PD 03-JUL-2003.
XX PF 07-MAY-2001; 2001US-00851410.
XX PR 17-JUN-1988; 88US-00208997.
XX PR 11-APR-1989; 89US-0036672.
XX PR 16-JUN-1989; 89US-00367486.
XX PR 13-OCT-1989; 89US-00420921.
XX PR 05-APR-1990; 90US-00505888.
XX PR 25-APR-1991; 91US-00681078.
XX PR 25-JUL-1994; 94US-00279823.
XX PR 03-AUG-1998; 98US-00128275.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam AW;
XX PI Fry KE;
XX DR WFI; 2003-810978/76.
XX DR N-PSDB; ADD24374.
XX PT New proteins derived from an enterically transmitted non-A/non-B (ET-
XX PT NANB) viral hepatitis agent, useful for diagnosing ET-NANB infection and
XX PT producing vaccines for immunizing an individual against ET-NANB hepatitis
XX PT viral agent.
XX PS Disclosure; SEQ ID NO 8; 81pp; English.
XX CC The invention relates to a protein derived from an enterically
XX CC transmitted non-A/non-B viral hepatitis agent. The viral proteins are
XX CC useful for detecting infection by the enterically transmitted non-A/non-B
XX CC (ET-NANB/HEV) hepatitis viral agent. The proteins are also useful for
XX CC producing vaccines for immunising an individual against HEV hepatitis
XX CC viral agent. The DNA fragments are useful for identifying and sequencing
XX CC the entire viral agent and for assaying the viral agent in an infected
XX CC sample. The present sequence is used in the exemplification of the
XX CC invention.
XX SK Sequence 660 AA;

Query Match 100.0%; Score 3408; DB 7; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.3e-263;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPRPILLLLMLPMLPAPPQPSGRRRGRSGSGGFWGDRVDSQFAIPYIHTN 60
Db 1 MRPRPILLLLMLPMLPAPPQPSGRRRGRSGSGGFWGDRVDSQFAIPYIHTN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQPAVASRRRPTTAGAAPTATAPAHDT 120
Db 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQPAVASRRRPTTAGAAPTATAPAHDT 120
QY 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRELVPNAVGGYAISISFWPQTITPTSDMNSITSDVRLVOPGI 240
Db 181 NYAQYRVARATIRYRELVPNAVGGYAISISFWPQTITPTSDMNSITSDVRLVOPGI 240
QY 241 ASELVIPSERLHYRNOGWSRVETSGVAEEBATSGLVWLCHIGSLVNSYNTNPTYGALL 300
Db 241 ASELVIPSERLHYRNOGWSRVETSGVAEEBATSGLVWLCHIGSLVNSYNTNPTYGALL 300
QY 301 DFALELEFRNLTPCNTNTRVSRYSSTARHLRRGADGTAEITTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFRNLTPCNTNTRVSRYSSTARHLRRGADGTAEITTAATRFMKDLYFTSTNG 360
QY 361 VGEIRGIALTLFNLADTLGGLPTELISSAGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIRGIALTLFNLADTLGGLPTELISSAGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIALPHDIDLGESRVVIQDYDNOHQEODRPTSPAPSRPFVLRANDVWLSLTAAEY 480
Db 421 QDKGIALPHDIDLGESRVVIQDYDNOHQEODRPTSPAPSRPFVLRANDVWLSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVLDGRPLSTIQOYSKTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVLDGRPLSTIQOYSKTFFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660
Db 601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660

RESULT 11
AAW76369
ID AAW76369 standard; protein; 560 AA.
XX AC AAW76369;
XX DT 03-DEC-1998 (first entry)
XX DE Hepatitis E virus hollow particle protein #2.
XX KW Hollow particle protein; virus; antibody; detection; immunoassay;
XX KW infection.
XX OS Hepatitis virus.
XX PN JF10234383-A.
XX PD 08-SEP-1998.
XX PF 28-FEB-1997; 97JP-00062445.
XX PR 28-FEB-1997; 97JP-00062445.
XX PA (DENK-) DENKA SEIKEN KK.

(KOKU-) KOKURITSU YOB0 EISEI KENKYUSHO.
WPI; 1998-535037/46.
N-PSDB; AAV61688.
Hepatitis E virus hollow particle poly-peptide(s) and nucleic acids
encoding it - useful for more accurate detection of HEV in samples, using
immuno-assays and nucleic acid hybridisation.
Claim 13; Page 24-26; 29pp; Japanese.
This sequence represents a Hepatitis E viral hollow particle protein.
This polypeptide can be used to raise antibodies to detect HEV infection
in samples, e.g. by immuno-assay based techniques, and the nucleic acid
can be used for the same in nucleic acid hybridisation assays. The
CC polypeptides and nucleic acids allow more accurate detection of HEV than
CC previously possible
XX
XX Sequence 660 AA;
Query Match 99.8%; Score 3401; DB 2; Length 660;
Best Local Similarity 99.8%; Pred. No. 1.9e-262;
Matches 659; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MRPRPILILLMLPMLPAPPQPSGRRRRSGSGGFWGDRVDSQPPAIPYIHPTN 60
Db 1 MRPRPILILLMLPMLPAPPQPSGRRRRSGSGGFWGDRVDSQPPAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSAMRDQORPAVASRRRPTTAGAAPTAVAPAHDT 120
Db 61 PFAPDVTAAAGAGPRVQPARPLGSAMRDQORPAVASRRRPTTAGAAPTAVAPAHDT 120
QY 121 PVPDVSRRGAILLRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEAS 180
Db 121 PVPDVSRRGAILLRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVNAVGGYAIISFWPQTTTTPTSDMNSITSTDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGGYAIISFWPQTTTTPTSDMNSITSTDVRLVQPGI 240
QY 241 ASELVIPERLHYRNQGWHSVTSQVAEEATSGLVMLCIHGSVNSYNTPTTGALGLL 300
Db 241 ASELVIPERLHYRNQGWHSVTSQVAEEATSGLVMLCIHGSVNSYNTPTTGALGLL 300
QY 301 DFALSEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELETTTAAATPMKDYFTSTNG 360
Db 301 DFALSEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELETTTAAATPMKDYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLTLSSAGGQIFYGRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLTLSSAGGQIFYGRPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGSRVVIQYDQNHQDRTPSPAPSRPESVLRANDVLSLTAAY 480
Db 421 QDKGIAIPHDIDLGSRVVIQYDQNHQDRTPSPAPSRPESVLRANDVLSLTAAY 480
QY 481 DQSTYGSSTGVPYVSDSVTLNVNATGAQVARSLDWTKVTLIDGRPLSTIQYKSTFFVLP 540
Db 481 DQSTYGSSTGVPYVSDSVTLNVNATGAQVARSLDWTKVTLIDGRPLSTIQYKSTFFVLP 540
QY 541 LRKLSFWEAGTTKAGYPYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKLSFWEAGTTKAGYPYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHASALLEDTLDYPARAHTFDDFCPCPLGLOGCAFQSTVAELQRLKMKVGTREL 660
Db 601 LAPHASALLEDTLDYPARAHTFDDFCPCPLGLOGCAFQSTVAELQRLKMKVGTREL 660
RESULT 12
AAW71210
ID AAW71210 standard; protein; 660 AA.
XX

AAW71210;
AC
XX
XX 25-MAR-2003 (revised)
DT 30-OCT-1998 (first entry)
XX
XX
XX Protein encoded by ORF 2 of the Burmese isolate of ET-NANB.
XX Enterically transmitted nonA/nonB hepatitis virus; identification; HEV;
XX ET-NANB; detection; vaccine.
XX
XX Hepatitis virus.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 496
FT /note= "not specified"
XX
XX
XX US5789559-A.
XX
XX 04-AUG-1998.
XX
XX 25-JUL-1994; 94US-00279823.
XX
XX 17-JUN-1988; 88US-00208997.
XX 11-APR-1989; 89US-00336672.
XX 19-JUN-1989; 89US-00367486.
XX 13-OCT-1989; 89US-00420921.
XX 05-APR-1990; 90US-00505888.
XX 05-APR-1991; 91US-00661078.
XX
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX
XX Bradley DW, Reyes GR, Krawczynski KZ, Tam A, Fry KE, Yarbrough PO;
XX WPI; 1998-446186/38.
XX
XX Hepatitis E virus DNA - useful for e.g. virus detection and viral protein
XX production.
XX
XX Disclosure; Col 63-68; 45pp; English.
XX
XX AAW71209-11 represent the proteins encoded by the open reading frames
XX (ORFs) of the DNA sequence of the Burmese isolate of an enterically
XX transmitted nonA/nonB viral hepatitis agent (ET-NANB). The nucleic acid
XX sequence may be used for identifying and sequencing the entire viral
XX agent (also referred to as HEV), detecting ET-NANB in infected samples,
XX e.g. by specific amplification of virus-derived DNA sequences and for
XX producing recombinant viral proteins for use in vaccines. (Updated on 25-
XX MAR-2003 to correct PF field.)
XX
XX Sequence 660 AA;
Query Match 99.8%; Score 3401; DB 2; Length 660;
Best Local Similarity 99.8%; Pred. No. 1.9e-262;
Matches 659; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MRPRPILILLMLPMLPAPPQPSGRRRRSGSGGFWGDRVDSQPPAIPYIHPTN 60
Db 1 MRPRPILILLMLPMLPAPPQPSGRRRRSGSGGFWGDRVDSQPPAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSAMRDQORPAVASRRRPTTAGAAPTAVAPAHDT 120
Db 61 PFAPDVTAAAGAGPRVQPARPLGSAMRDQORPAVASRRRPTTAGAAPTAVAPAHDT 120
QY 121 PVPDVSRRGAILLRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEAS 180
Db 121 PVPDVSRRGAILLRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVNAVGGYAIISFWPQTTTTPTSDMNSITSTDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGGYAIISFWPQTTTTPTSDMNSITSTDVRLVQPGI 240
QY 241 ASELVIPERLHYRNQGWHSVTSQVAEEATSGLVMLCIHGSVNSYNTPTTGALGLL 300
Db 241 ASELVIPERLHYRNQGWHSVTSQVAEEATSGLVMLCIHGSVNSYNTPTTGALGLL 300

Db	181	NYAQVRVARATIRYRLVFNAGGVAISLFWPQTTTTFTSVDMNSITSTDVRLVQPGL	240
Qy	241	ASELVIPSERLHYRNQGRSVETSGVAEEEAISGLVMLCIHGSLVNSYTNTPYTGAIGLL	300
Db	241	ASELVIPSERLHYRNQGRSVETSGVAEEEAISGLVMLCIHGSPVNSYTNTPYTGAIGLL	300
Qy	301	DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTAATRFMKDLYFTSTNG	360
Db	301	DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTAATRFMKDLYFTSTNG	360
Qy	361	VGEIGRGIALTLPNLADTLLGGIPTLTISSAGGQLFYSRPVVSANGEPVVKLYTSVENAQ	420
Db	361	VGEIGRGIALTLPNLADTLLGGIPTLTISSAGGQLFYSRPVVSANGEPVVKLYTSVENAQ	420
Qy	421	QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVLMLSLTAAY	480
Db	421	QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVLMLSLTAAY	480
Qy	481	DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLDCRPLSTIQOYSKTFVLP	540
Db	481	DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLDCRPLSTIQOYSKTFVLP	540
Qy	541	LRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTVYTTSLGAGPVSISAVAV	600
Db	541	LRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTVYTTSLGAGPVSISAVAV	600
Qy	601	LAPHSALELLEDLDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGKTREL	660
Db	601	LAPHSVLALLEDTMDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGKTREL	660

Search completed: August 16, 2004, 13:33:47
Job time : 59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 13:28:02 ; Search time 20 Seconds
(without alignments)
3174.320 Million cell updates/sec

Title: US-09-851-410A-8
Perfect score: 3408
Sequence: 1 MRPRPILLLLMLFPLMLPAP.....QSTVAELQRLKMKVGKTRLE 660
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3408	100.0	660	1 VHWWH2	structural protein
2	3194.5	93.7	659	2 B44212	structural protein
3	143	4.2	1041	2 S55862	probable membrane
4	139.5	4.1	5188	2 B85547	probable RTX famil
5	139.5	4.1	5291	2 F90696	hypothetical prote
6	137	4.0	3570	2 T45025	mucin MUC5B, trach
7	135.5	4.0	1324	2 S52863	DNA-binding protei
8	135.5	4.0	3282	2 E82750	hemagglutinin-like
9	135.5	4.0	3442	2 E82589	hemagglutinin-like
10	135.5	4.0	3455	2 B82519	hemagglutinin-like
11	133	3.9	796	2 T21460	hypothetical prote
12	133	3.9	1802	2 S69703	HKR1 protein precu
13	132.5	3.9	583	2 S67571	hypothetical prote
14	130.5	3.8	1104	2 S93110	hypothetical prote
15	130.5	3.8	1161	2 S57180	probable membrane
16	130.5	3.8	1537	2 S33465	probable membrane
17	130.5	3.8	3013	2 A80480	flocculation prote
18	129.5	3.8	635	2 F75477	probable invasin y
19	129.5	3.8	1367	1 S48478	hypothetical prote
20	129.5	3.8	2130	2 A80821	glucan 1,4-alpha-g
21	129	3.8	918	2 T02759	probable exported
22	129	3.8	1015	2 J05062	hypothetical prote
23	128.5	3.8	1367	2 S51959	phogrin precursor
24	128	3.8	767	1 J00474	hypothetical prote
25	128	3.8	778	1 ALBYG	glucan 1,4-alpha-g
26	128	3.8	832	2 B87673	glucan 1,4-alpha-g
27	127.5	3.7	518	1 P4WL47	ABC transporter, H
28	127.5	3.7	800	2 T02852	L2 protein - human
29	127.5	3.7	860	2 T14650	probable membrane
					tail fiber protein

30	127.5	3.7	978	2 T14968	phage lambda-relat
31	127	3.7	1275	2 T33369	hypothetical prote
32	127	3.7	7463	2 T36248	CDA peptide synthe
33	126.5	3.7	351	1 VCETC	coat protein - tur
34	126.5	3.7	948	2 T11678	hypothetical prote
35	126.5	3.7	1306	2 S25370	MSB2 protein - yea
36	126	3.7	1037	2 T13350	transcription fact
37	125.5	3.7	725	2 A41258	a-agglutinin core
38	125.5	3.7	1879	2 S74915	extracellular nucl
39	125.5	3.7	1952	2 T48814	hypothetical prote
40	124	3.6	963	2 T40290	hypothetical prote
41	124	3.6	1131	2 T41144	hypothetical serin
42	124	3.6	1483	2 C97012	probably celluloso
43	123.5	3.6	797	1 VGBEX1	glycoprotein X pre
44	123	3.6	2322	2 T34434	hypothetical prote
45	122.5	3.6	1324	2 T14070	peptide synthetase

ALIGNMENTS

RESULT 1

VHWWH2
structural protein 2 precursor - hepatitis E virus (strain Burma)
C:Species: hepatitis E virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 23-Jul-1999
C:Accession: C40778
R:Tam, A.W.; Smith, M.M.; Guerra, M.E.; Huang, C.C.; Bradley, D.W.; Fry, K.E.; Reyes, G.
Virology 185, 120-131, 1991
A:Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length vi
A:Reference number: A40778; MUID:92024067; PMID:1926770
A:Accession: C40778
A:Molecule type: genomic RNA
A:Residues: 1-660 <TAM>
A:Cross-references: GB:M73218; NID:G330023; PID:AAA45736.1; PID:G330026
A:Note: the authors translated the codon CGC for residue 2 as Ala
C:Superfamily: hepatitis E virus structural protein 2
C:Keywords: structural protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-660/Product: structural protein 2 #status predicted <SP2>

Query Match 100.0%; Score 3408; DB 1; Length 660;
Best Local Similarity 100.0%; Pred. No. 8.2e-208; Indels 0; Gaps 0;
Matches 660; Conservative 0; Mismatches 0;

QY	1	MRPRPILLLLMLFPLMLPAPPPQPSGRRRSGSGGGGFWGDRVDSQPPFAIPYIHPTN	60
DB	1	MRPRPILLLLMLFPLMLPAPPPQPSGRRRSGSGGGGFWGDRVDSQPPFAIPYIHPTN	60
QY	61	PPAPDVTAAGAGPRVQPARPLGSAWDDQAPAVASRRRPTTAGAAPLTAVAPADTP	120
DB	61	PPAPDVTAAGAGPRVQPARPLGSAWDDQAPAVASRRRPTTAGAAPLTAVAPADTP	120
QY	121	PVPDVDSRGAILRROYNLSTPLTSSVATGNLVLYAAPLPLPLQDGTTHINATAS	180
DB	121	PVPDVDSRGAILRROYNLSTPLTSSVATGNLVLYAAPLPLPLQDGTTHINATAS	180
QY	181	NYAQVRVARATIRVPLVPNAVGGVAISISFWPQTITPTSDVMNSIITSDVRILVQPGI	240
DB	181	NYAQVRVARATIRVPLVPNAVGGVAISISFWPQTITPTSDVMNSIITSDVRILVQPGI	240
QY	241	ASELVIPISERLHYRNOQRVSTSGVAEEBETAGLWLCIHGSLVNSVNTPTYGALGILL	300
DB	241	ASELVIPISERLHYRNOQRVSTSGVAEEBETAGLWLCIHGSLVNSVNTPTYGALGILL	300
QY	301	DFALELFRNLTGNTNTRVSRYSSTAHRLRRGADGTAEITTTAATFRMKDLYFTSTNG	360
DB	301	DFALELFRNLTGNTNTRVSRYSSTAHRLRRGADGTAEITTTAATFRMKDLYFTSTNG	360
QY	361	VGEIGRGIALTLFNLAIDLTLGLLPTELISAGGOLFYSRPVVSANGEPVTKLYTSVENAQ	420
DB	361	VGEIGRGIALTLFNLAIDLTLGLLPTELISAGGOLFYSRPVVSANGEPVTKLYTSVENAQ	420

QY 421 QDKGIAIPHDIDLGSRVVIQYDNOHQEODRTPSPAPSRPSVLRANDVLMLSLTAAY 480
DB 421 QDKGIAIPHDIDLGSRVVIQYDNOHQEODRTPSPAPSRPSVLRANDVLMLSLTAAY 480
QY 481 DOSTYGSSTGPPVVSQSVTLNVNATCAQAVARSLDWTVKTLDRPLSTTQQYSKTFVFLP 540
DB 481 DOSTYGSSTGPPVVSQSVTLNVNATCAQAVARSLDWTVKTLDRPLSTTQQYSKTFVFLP 540
QY 541 LRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRAVAISTYTTSLGAGPVSISAVAV 600
DB 541 LRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRAVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSLALLEDTLDYPARAHTFDDFCPCPCPLGLQGCFAQSVAEQLRLKMKVGTREL 660
DB 601 LAPHSLALLEDTLDYPARAHTFDDFCPCPCPLGLQGCFAQSVAEQLRLKMKVGTREL 660
RESULT 2
B44212
structural protein 2 precursor - hepatitis E virus (strain Mexico)
C:Species: hepatitis E virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: B44212
R:Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.
Virology 191, 550-558, 1992
A:Title: Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HE
A:Reference number: A44212; MUID:93079857; PMID:1448913
A:Accession: B44212
A:Molecule type: Genomic RNA
A:Residues: 1-659 <HUA>
A:Cross-references: GB:M74506; NID:G330017; PIDN:AAA45732.1; PID:G330020
C:Superfamily: hepatitis E virus structural protein 2
C:Keywords: structural protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-659/Product: structural protein 2 #status predicted <SP2>
Query Match 93.78; Score 3194.5; DB 1; Length 659;
Best Local Similarity 93.08; Pred. No. 2.6e-194;
Matches 614; Conservative 21; Mismatches 24; Indels 1; Gaps 1;
QY 1 MRPRPILLMLFPLPAPPGQSGRRGRSGSGGFGWGDVDSQPPFAIPYIHTN 60
DB 1 MRPRPILLMLFLLPMLPAPTQSGRRGRSGGTGGFGWGDVDSQPPFAIPYIHTN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSAMRDQAPAVASRRRPTTAGAAPLTAVAPADTP 120
DB 61 PFAPDVAAASGGPRLQPARPLGSTWRDQAPESASRRRPATAGAAALTAVAPADHTS 120
QY 121 PVPDVSRGATLRQYNLSTSPLTSSVATGNLVLYAAPLSPLPLQDGTNTHIMATEAS 180
DB 121 PVPDVSRGATLRQYNLSTSPLTSSVASGNLVLYAAPLNPPPLQDGTNTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVPNAVGGYVAISISFPQTPTTPTSDVMNSITSDVRLVQPGI 240
DB 181 NYAQYVARATIRYRPLVPNAVGGYVAISISFPQTPTTPTSDVMNSITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNQWRVETSGVAEEATSGVLMCLIHGSLVNSVNTPTPYTGALGL 300
DB 241 ASELVIPSERLHYRNQWRVETSGVAEEATSGVLMCLIHGSPVNSYNTPTPYTGALGL 300
QY 301 DFALELEFRNLTPCNTNTRVSRYSSTARHRRRGADGTAEITTTAAATFRMKDLYFTSTNG 360
DB 301 DFALELEFRNLTPCNTNTRVSRYSSTARHS-ARGADGTAEITTTAAATFRMKDLFTGLNG 359
QY 361 VGEIGRGIALTFNLADTLGLLPTLTISSAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
DB 360 VGEVGRGIALTLNLADTLGLLPTLTISSAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 419
QY 421 QDKGIAIPHDIDLGSRVVIQYDNOHQEODRTPSPAPSRPSVLRANDVLMLSLTAAY 480
DB 420 QDKGVAIPHDIDLGSRVVIQYDNOHQEODRTPSPAPSRPSVLRANDVLMLSLTAAY 479
QY 481 DOSTYGSSTGPPVVSQSVTLNVNATCAQAVARSLDWTVKTLDRPLSTTQQYSKTFVFLP 540

DB 480 DOSTYGSSTGPPVVSQSVTLNVNATCAQAVARSLDWSKVTLDRPLPTVEQYSKTFVFLP 539
QY 541 LRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRAVAISTYTTSLGAGPVSISAVAV 600
DB 540 LRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRAVAISTYTTSLGAGPVSISAAAV 599
QY 601 LAPHSLALLEDTLDYPARAHTFDDFCPCPCPLGLQGCFAQSVAEQLRLKMKVGTREL 660
DB 600 LAPRSALALLEDTDYPGRAHTFDDFCPCPCRALGLQGCFAQSVAEQLRLKMKVGTREL 659
RESULT 3
S55862
Probable membrane protein YNL327w - Yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein N0320
C:Species: Saccharomyces cerevisiae
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
C:Accession: S55862; S63308; S51288
R:Maftahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
Yeast 11, 567-572, 1995
A:Title: Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV identifies the
A:Reference number: S55859; MUID:95373280; PMID:7645347
A:Accession: S55862
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1041 <MAF>
A:Cross-references: EMBL:Z46259; NID:G633655; PIDN:CAA86371.1; PID:G633659
R:Maftahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63287
A:Accession: S63308
A:Molecule type: DNA
A:Residues: 1-1041 <MAW>
A:Cross-references: EMBL:271603; NID:G1302444; PID:e239572; PID:G1302445; MIPS:YNL327w
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:EGT2
A:Cross-references: SGD:S0005271; MIPS:YNL327w
A:Map position: 14L
C:Keywords: transmembrane protein
F:4-20/Domain: transmembrane #status predicted <TM1>
F:1025-1041/Domain: transmembrane #status predicted <TM2>
Query Match 4.2%; Score 143; DB 2; Length 1041;
Best Local Similarity 20.48; Pred. No. 0.48;
Matches 144; Conservative 94; Mismatches 238; Indels 230; Gaps 31;
QY 48 SQPFAIPYIHTNPFAPDVTAAGAGPRVQPARPLGSAMRDQAPAVASRRRPTT--- 104
DB 340 SQSSAIPV-----AVTASLSSG-----ILSSTVDGASTSADASMSAVSVSS 382
QY 105 ----AGAAPLTAVAPADHTPPVPDVSRGATLRQYNLSTSPLTSSVATGNLVLYAAPL 160
DB 383 SSEQASSSISLSAPSSN-----STFTTPSSSLGATETYSIISASAI 425
QY 161 SPLPLQDGTNTHIMATEASNYAQYRVAR---ATIRYRPLVPNAVGGYVAISISFPQTPT 217
DB 426 SVTQASVIDNSTTTAVTQSTSTIAVSSAEKLSSTLSY-----TSNVTISVSSATQHTT 478
QY 218 TPTSDVMNSITSDVRLVQPGIASSELVIPSERLHYRNQWRVETSGVAEEATSGVLM 277
DB 479 TPSVVS-NSITLSSSVL-----TCGALGLDFALELEFRNLTPGNTNTRVSRYSSTAR 511
QY 278 LCIHGLSVNSVNTPY-----TGALGLDFALELEFRNLTPGNTNTRVSRYSSTAR 328
DB 512 ----ASSASOSTNPYPVNSNSTSSATQIATAPPAINTIGTSSISSTNT--SVSSYIT- 564
QY 329 HRLRRG-----ADGTAELTTTAAATFRMKDL-----YFTSTNGVGEIRG--- 367
DB 565 SSLSSGPPVSNATVAGSGVILTTTTSQAQLTEIGSLIPISITTTTSTGTDKTCGNKVA 624
QY 368 ----IALTLFNLADTLGLLTELISIA-----GGQL---FYSRPVVSANGE 407

```
Db 625 SSTEIAQSVNNSSLSVSTINTNAATAAANARNATEATHHSGSLQSPSYHSSLLSSTD 684
QY 408 PTVKLVTSVENAQQDGIAPHDIDIGESRVLQDYDNQHEQDRPTSPAPSPFVSLRA 467
Db 685 --TKVTATSTSRGSSSLAFTTGLNQSVVTGTD-----KSDTYSVIS 727
QY 468 NDVLMLSLAAEYDQ----STY-----GSSTGPVYVSDSVTLNVNATG-----AQAAR 512
Db 728 TE-----SAQVTEYDSLPLSTLKPTVVTGTSRNSTSMVSTKLTATATDKGDAYSVIS 783
QY 513 SLQWTKVTLQDGR--PLSTIQ-----QYSKTFVFLPLRGX 544
Db 784 STQSAQVTEYGMPLSTLETPVIMSTDESQVFTLTCTESQATEYGLIPISLTDGS 843
QY 545 LSPWEAG-----TTKAGYPNYNTASDQLLVENAAHRAVIAST-----YTSLG 589
Db 844 VIYFTGESVVGYSYTVGAAQVIAQHTSL---VPVSTIKGSKTSLSTESVWAGYSTTVG 900
QY 590 AG-----PVSI---SAVAVLAPHASALALLEDTLDYPARA-HT 622
Db 901 AAQYAQHTSIVPVSTIKGSKTSLSTESVVGYSYTVVDSQAQAEHT 946

RESULT 4
probable RTX family exoprotein [imported] - Escherichia coli (strain O157:H7, substrain
B85547
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: B85547
R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85547
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5188 <STO>
A:Cross-references: GB:AE005174; NID:G12513368; PIDN:AA054838.1; GSPDB:GN00145; UWGP:206
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z0615

Query Match 4.1%; Score 139.5; DB 2; Length 5188;
Best Local Similarity 22.3%; Pred. No. 8;
Matches 131; Conservative 70; Mismatches 235; Indels 151; Gaps 28;

QY 60 NPFAPDVTAAAGAGPRVQPARPLGSW-----RDOAQRPAVASRRRPTTAGAA 108
Db 2478 NSVAYSATVQADGWSVGIIPAANV-SAWPAGPLTVEVDGQSSANNPVS--HPFT--V 2531
QY 109 PLTAVAPADHTPPVDPVDSRGAILRRQYNLSTPSLTSSVATGTLNLVLAAPLSPLPLQD 168
Db 2532 DLTAVAISINTVASDDVINA--BKGNTLTLGSGTSGIESQTVT-----VTF 2577
QY 169 GNTHIMATEASNYAQYVARATIRYP-----LYPNAVGGYALISIFWPQTTPPT 220
Db 2578 GCKTITASVAANGSVNVVPAADLATLPEGAANVQASVSSAGNSASATHATYSDASAPT 2637
QY 221 SVMNMSITSDVRILVQPGIASSELVIPSERLHYRQGRSVETSGVABEEATSGLVMLCI 280
Db 2638 -LTINTIASDDILNAAEAG--SPLTI-----SGTSTAE-TGQIVTVTL 2676
QY 281 HGSLVNSTNTPY-----TGALGLDPALELEFRNLTPGNTNTRVSRYSSTARH 330
Db 2677 NGA---TYTGTVQADGWSVSVPTSALGALNAS-----NYTVSATVNDKAGNPGSASHN 2727
QY 331 LRRGADGTA---ELTTTAAATRFMKD-----LYFTSTNGVGEIGRGIALTLNLAADTLTG 381
Db 2728 L--AVDTTAPVLTINTVAGDDIINDAEHAQALVISGTSGGEXGDVSWVL----- 2776
QY 382 GLPTLISAGGQLFYSPVPSVANGEPTVKLYTSVENAQDQKIAIP-HDID-LGESRV 439
Db 331 LRRGADGTA---ELTTTAAATRFMKD-----LYFTSTNGVGEIGRGIALTLNLAADTLTG 381
QY 440 IODVNDQHEQDRPTSPAPSPFVSLRVANDVLWLSLTAEE-----YDQSTYGSSTG 490
Db 2810 I---TASVSDRAGNSDDASRTVTVLSAPVISINTIAGDDVINATEKGSLLALSGTSDQ 2865
QY 491 PVIYSDSVTLNVNATGAQAVARSLDWTKVTLQDGRPLSTIQ--YSKTFFVLPLRGKLSF- 547
```

```
Db 2777 -----NG-----KTYTTTLDASGNWSGVPAADVLTALGSAQT 2809
QY 440 IQDYDNQHEQDRPTSPAPSPFVSLRVANDVLWLSLTAAB-----YDQSTYGSSTG 490
Db 2810 I---TASVSDRAGNSDDASRTVTVLSAPVISINTIAGDDVINATEKGSLLALSGTSDQ 2865
QY 491 PVIYSDSVTLNVNATGAQAVARSLDWTKVTLQDGRPLSTIQ--YSKTFFVLPLRGKLSF- 547
Db 2866 PAGTALTIVTL-NGQNYSATTTDASGNWS-VTVPASAVSALGEATYSVTASVTNAQGNSSTA 2923
QY 548 --WEAGTTKAGYPNYNTASDQLLVENAAHRAVIASTVTTSLGAG 591
Db 2924 SHNVQNTALPG--ITINPVATDDIINASEASQAQTISQVTCGAAG 2968

RESULT 5
F90696
hypothetical protein ECs0542 [imported] - Escherichia coli (strain O157:H7, substrain RI
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: F90696
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F90696
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5291 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA033965.1; PID:G13360000; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs0542

Query Match 4.1%; Score 139.5; DB 2; Length 5291;
Best Local Similarity 22.3%; Pred. No. 8.2;
Matches 131; Conservative 70; Mismatches 235; Indels 151; Gaps 28;

QY 60 NPFAPDVTAAAGAGPRVQPARPLGSW-----RDOAQRPAVASRRRPTTAGAA 108
Db 2478 NSVAYSATVQADGWSVGIIPAANV-SAWPAGPLTVEVDGQSSANNPVS--HPFT--V 2531
QY 109 PLTAVAPADHTPPVDPVDSRGAILRRQYNLSTPSLTSSVATGTLNLVLAAPLSPLPLQD 168
Db 2532 DLTAVAISINTVASDDVINA--BKGNTLTLGSGTSGIESQTVT-----VTF 2577
QY 169 GNTHIMATEASNYAQYVARATIRYP-----LYPNAVGGYALISIFWPQTTPPT 220
Db 2578 GCKTITASVAANGSVNVVPAADLATLPEGAANVQASVSSAGNSASATHATYSDASAPT 2637
QY 221 SVMNMSITSDVRILVQPGIASSELVIPSERLHYRQGRSVETSGVABEEATSGLVMLCI 280
Db 2638 -LTINTIASDDILNAAEAG--SPLTI-----SGTSTAE-TGQIVTVTL 2676
QY 281 HGSLVNSTNTPY-----TGALGLDPALELEFRNLTPGNTNTRVSRYSSTARH 330
Db 2677 NGA---TYTGTVQADGWSVSVPTSALGALNAS-----NYTVSATVNDKAGNPGSASHN 2727
QY 331 LRRGADGTA---ELTTTAAATRFMKD-----LYFTSTNGVGEIGRGIALTLNLAADTLTG 381
Db 2728 L--AVDTTAPVLTINTVAGDDIINDAEHAQALVISGTSGGEXGDVSWVL----- 2776
QY 382 GLPTLISAGGQLFYSPVPSVANGEPTVKLYTSVENAQDQKIAIP-HDID-LGESRV 439
Db 2777 -----NG-----KTYTTTLDASGNWSGVPAADVLTALGSAQT 2809
QY 440 IODVNDQHEQDRPTSPAPSPFVSLRVANDVLWLSLTAEE-----YDQSTYGSSTG 490
Db 2810 I---TASVSDRAGNSDDASRTVTVLSAPVISINTIAGDDVINATEKGSLLALSGTSDQ 2865
QY 491 PVIYSDSVTLNVNATGAQAVARSLDWTKVTLQDGRPLSTIQ--YSKTFFVLPLRGKLSF- 547
```

Db 2866 PAGTALTIVTL-NGQNSATTDASGNWS-VTVFASVASLGEATYSYTSVTAQNSSTA 2923
Qy 548 ---WEAGTTKAGYPYNYNTTASDQLAVENAAGHRVAISYTTTSLGAG 591
Db 2924 SHNVQVNTALPG--ITINPVATDDIINASEAGSAQTISQVTTGAAAG 2968

RESULT 6
T45025
mucin MUC5B, tracheobronchial [imported] - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T45025
R:Desseyn, J.L.; Guyonnet-Duperrat, V.; Porchet, N.; Aubert, J.P.; Laine, A.
J. Biol. Chem. 272, 3168-3178, 1997
A:Title: Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alternat
A:Reference number: Z22899; MUID:97166151; PMID:9013550
A:Accession: T45025
A:Status: preliminary;
A:Molecule type: DNA
A:Residues: 1-3570 <DES>
A:Cross-references: EMBL:Z72496; NID:g1834502; PIDN:CAA96577.1; PID:g1834503
A:Experimental source: placenta
C:Genetics:
A:Gene: MUC5B

Query Match 4.0%; Score 137; DB 2; Length 3570;
Best Local Similarity 20.1%; Pred. No. 6.7;
Matches 143; Conservative 72; Mismatches 257; Indels 240; Gaps 33;

Qy 52 AIPYIH--PNNPAPDVTAAGAGPVQRPAPRLGSAWRDQAPPAVASRRRTTGAAP 109
Db 1867 ALPALRSTATTATTSVTAI-----PSSSLGTAW-----TRLSTQTTPTAT 1907
Qy 110 LTAVAPAHDTPPVPDVSRCAILLRQYNLSTPLTSSVATGTLNLVLAAPLSPLLPDQG 169
Db 1908 MSTATPS-STPE-----TVHTSVLTATTTRGVSATPSS-----TPG 1946
Qy 170 T--NTHIMATEAGNYAQYVARATIRYRPLVFNAGVYASISFWPQTTTTPT 220
Db 1947 TAHTTKVPTTTTGT-----ATPSSSPGALTTPVWISTTTTTPTTRGSTVT 1993
Qy 221 --SVDMNSITSDVRILVQGIASELVIPIERLHYRQGRSVETSGVAEEATSLGLVL 278
Db 1994 PSSIPGTHATVLTITTTTVAATGSMATPSS-----STQTSCTPPSLTTATIT 2043
Qy 279 CIHGSLVN-----SYTNTPYTGALG-----LLDFALELF 308
Db 2044 AT-GSTNPSSTPGTPIPPVLITATTATPAATSSVTP-SSALGTHTPPVNTATTGG 2101
Qy 309 RNLTGNTNTRVRSYSTARHLRRLRGADGTAELT-----TTAAT----- 347
Db 2102 RSLPPSPHTVPTAWTSAT-----SGILGTHITEPSTGTSHTPAATTGTTQPSALSS 2156
Qy 348 -----RPMKDLFTSTNGVGEICRGIALTLFNL-----ADTLGLGLPTE-----LISSAG 392
Db 2157 PHPSSRTSPSPSGTTTPGHT-RGTSRTTATATPSTRKTSLLPSSPTSAPITVVITG 2215
Qy 393 GQ-----LFYSRPVVSANGEPTVKLYTSVE-----NAQQDKGI----- 425
Db 2216 CEPQCAWSEWLDYFYPGPGSG-GDFETYSNIRAAGGAVCEQPLGLECRAQAQGVPLRE 2274
Qy 426 ---AIPHDIDL-----ESRVVIQDYNQHQDRP-----TPSPAP 458
Db 2275 LGQVVECSLDFGLVCRNREQVKFCMCFNVEIRVFCNCG--HCFSTPATSSATPSSTP 2332
Qy 459 SRPFSVLRANDVLMLSLTAAYD-----QSTYGSSTGPVYVSDSVLNVNATGAQAV 510
Db 2333 GTTWILTK-----LTTATTTESTGSTATPSSTQGGPAGTPHVSHTATTFTV-TSSKAT 2385
Qy 511 ARSLDWTKVTLDRPLSTIQOYSKTFVFLPLRGLKLSFWEAGTTKAGYPYNYNTTASDQL 570

Db 2386 PFSSPGTATALPALRSTATTATTSATPSSTSLGTTW-TRLSTQTTTPMATMSTATPSST 2444
Qy 571 VENAGHRVAISYTTTSLGAGPVVISAVAVLAPHLSALALLEDTLDVYPARAHT 622
Db 2445 PETVHTSTVLTTATTATTGATG-----SVATPSST-----PGTAHT 2479

RESULT 7
S52863
DNA-binding protein R kappa B - human
C:Species: Homo sapiens (man)
C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 01-Dec-2000
C:Accession: S52863; A45580
R:Nieters, A.; Boumeester, T.; Scheidereit, C.
submitted to the EMBL Data Library, August 1994
A:Reference number: S52863
A:Accession: S52863
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1324 <NIE>
A:Cross-references: EMBL:X80878; NID:g695578; PID:g695579
R:Adams, B.S.; Leung, K.Y.; Hanley, E.W.; Nabel, G.J.
New Biol. 3, 1063-1073, 1991
A:Title: Cloning of R kappa B, a novel DNA-binding protein that recognizes the interleuk
A:Reference number: A45580; MUID:92135142; PMID:1777480
A:Accession: A45580
A:Molecule type: mRNA
A:Residues: 313-1292, 'ISNREQLIAPP', 1304, 'QSPLLL', <ADA>
A:Cross-references: GB:U08191; GB:S79520; NID:g476273; PIDN:AAA1781.1; PID:g476274
A:Note: sequence extracted from NCBI backbone (NCBIN:79520)
C:Keywords: DNA binding; transcription factor

Query Match 4.0%; Score 135.5; DB 2; Length 1324;
Best Local Similarity 19.0%; Pred. No. 2;
Matches 131; Conservative 95; Mismatches 232; Indels 231; Gaps 29;

Qy 68 AAAGAGPRVRQARP-----LGSARWDQAPPAVASRRRT-----TA 105
Db 693 AAARKALQKPKPKPKSVKSSKSSIKVLSSGPSEQSOMSLSDSMPTTPVTPVPTT 752
Qy 106 GAAPLTAVAPAHDTPPVPDVSRCAILLRQYNLSTPLTSSVATGTLNLVLAAPLSPL 164
Db 753 PALPALPISP-----PPVSAVNKSG-----PSTVSEPAKSSGVLVSSSTPHLG 798
Qy 165 -----PLQDGTNTHIMATEASNVAAQYR-VARATIRYRPLVFNAGVYASIS-- 209
Db 799 TMLSPASSQTAPSSQAARVSHSGSAGLSQVRVVAQPSL---PAVQQSGGPAQTLPQM 855
Qy 210 -----SFWPQT-----TTPTSVDMNSITSDVRILVQGIASELV 245
Db 856 PAGQIRVPATATQTKVVPQTMATVPVKAQTAAATVQRPQGQTLTSLPATASPV 915
Qy 246 IPIERLHYRQGRSVETSGVAEEATSLGLVLCIHGSLVNSYNTTPTGALG----- 299
Db 916 KP-----ATSSPGTSAPSAATAAVIQNVTGQ--NIKQVAITGOLGVKVPOTGN 961
Qy 300 -----LDFALE-LEPRNLTPGNTNTRVRSYSTARHLRRLRGADGTAELTTAAT--RFM 350
Db 962 SIPLTATNFRIGQKDVLRPLPSSITT-----DAKQIVLRIPTDMMATLAKSQVTTVKLT 1016
Qy 351 KDLYFTSTNGVBRIGRIALTFLFNADTLGLGPTLGLISSAGGOLFYSRPVVSANGPTV 410
Db 1017 QDLFGTGCN--TTGKGISATLHVTN-----PVAADSPAKASSASAFSSTPTGTTV 1067
Qy 411 KLYTSVENAQQDK-----GTAIHDIIDLGESRVVIQDYNQHQDRPTSPAPSRP 461
Db 1068 KVTPLDKPTTASSAFRLMPALGVSV-----ADQKSTVASSAKP 1109
Qy 462 FSVLRANDVLMWL-----SLTAEYDQ--STYGSSTGPVYVSDSVTL-----VN 502
Db 1110 AATIRIVQGLGVMPKPAQGITVATHAKQGSVASGSGTVHTS-AVSLPSMNAAVSKTVA 1168
Qy 503 VATGAQAVARSLDW-----TKVTLDRPLSTIQOYSKTFV 538

```

Db      1169 VASGAASSTPISGTGAPTRQVFPVSTVSTQAGKLPTRITV---PLSVISQ----- 1218
QY      539 LPLRGK-----LSFWAGT---TKAGYPVNYNTASDQLLVE 572
Db      1219 -PMKGSVVTAPIKGNLNSGLGRNIILTMPAGTKLIAGNKPVSFLTAAQLOOLOQ 1277
QY      573 NAAGHRVAISTYTTTS-LGAGPVSISAVAV 600
Db      1278 QCATQVRIOTVPASHLQQGTASGSSKAV 1306

RESULT 8
hemagglutinin-like secreted protein XF0889 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: E82750
R:anonymus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82750
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3282 <SIM>
A:Cross-References: GB:AE003928; GB:AE003849; NID:g9105798; PIDN:AAF93699.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sancelli, R.V.; Sawaasak
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0889

Query Match      4.0%; Score 135.5; DB 2; Length 3282;
Best Local Similarity 21.9%; Pred. No. 7.4;
Matches 135; Conservative 64; Mismatches 216; Indels 201; Gaps 27;

QY      97 ASRRRPTTAGAAPTAVAPADTP-----PVP-----DVDSRGAIL 132
Db      429 ASPADPT-----PVASVPPTATPTSTGTPIPSPQTPAVLAAGRLTLENDIDNRGGHI 483
QY      133 RQYNLSTSPITSSVATGNLVLYAAPLSPL-LPLQDGTNTH---IMATEA-----SN 181
Db      484 TAGGAIDAILTLNDRGT-----AALNRLTQOQLDNOHQILTATEATHHTLNN 537
QY      182 YQYRVARATIRYPLVNAVGGVAISIFWPQTTTPTSDVMSITSDVRLVQPGIA 241
Db      538 AAGQLHVNGTLDLTAQHLNNAAGHLLTTG--PQSATL-TIADLLDNTSGTL-----ASA 588
QY      242 SELVIPERLHYRNGQWRSVETSGVAEEBATSGLVMLCIHGLSVNSYNTPTYGALGLLD 301
Db      589 GSLFLTAATL-----DITDGTVOSGGPLHMDAATLTAHRTGLTTSQD 630
QY      302 FALELEFRNLTPGNNTNRSYSSTAR-----HRLRGA----- 335
Db      631 -----TLTLTGHTDLSHATTAAQOITLHTGTLTAGGHTAYGTHTLQLEARTLTD 682
QY      336 -----DGTAEITTTAATPMKDLYFTSNGVGEIGRGIALTLFNLADTLGLGLPTPEL 387
Db      683 NTGGTIATNGTLDLHTAALDNTGGTLHSTAT---GPNRLDITDTLTNTAGHLLNGPTTL 739

```

```

QY      388 I----SSAGQQLFYSRPVV-----SANGEPVTKLYTSVENAQDQKGIAPH 429
Db      740 TTGTWTTNGHQHQTITPATLHATLNDNRGILHTATGPLDRLTRITGTLNN--QDNGLLTTT 797
QY      430 D-IDLGESRVVIQDYDNQH-EQDRPTPSPAP-----SRPFSVLRANDVWL----- 473
Db      798 DALTLTAASLL-----NQHTLDAAGPAHLTTLGLDNTAGLLOTAHTLWLSAGLTNRS 852
QY      474 -SLTAAEYDQSTYGS-----TGPV-----YVSDSVTL 500
Db      853 GTLTASQLTLDTCQHTLNDNRGRGLTTCGNLTLSHSGPLDNTAGLLOTAALTTDTGATL 912
QY      501 VNVATGAQAVARSLDWTKVTLGRPLSTIQOYSKIFFVLPLRGLKLSFWAGTTKAGYPYN 560
Db      913 TNRDGGALLAATALDLDLTATLNDNRG-GTIDSQTAT-----HLHTTLDNTTAGHSS 963
QY      561 YNTTASDQLLVENAAAG 576
Db      964 SGTLOIDGTTLNTGG 979

RESULT 9
hemagglutinin-like secreted protein XF2196 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: E82589
R:anonymus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82589
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3442 <SIM>
A:Cross-References: GB:AE004032; GB:AE003849; NID:g9107324; PIDN:AAF94995.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sancelli, R.V.; Sawaasak
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2196

Query Match      4.0%; Score 135.5; DB 2; Length 3442;
Best Local Similarity 21.9%; Pred. No. 7.9;
Matches 135; Conservative 64; Mismatches 216; Indels 201; Gaps 27;

QY      97 ASRRRPTTAGAAPTAVAPADTP-----PVP-----DVDSRGAIL 132
Db      429 ASPADPT-----PVASVPPTATPTSTGTPIPSPQTPAVLAAGRLTLENDIDNRGGHI 483
QY      133 RQYNLSTSPITSSVATGNLVLYAAPLSPL-LPLQDGTNTH---IMATEA-----SN 181
Db      484 TAGGAIDAILTLNDRGT-----AALNRLTQOQLDNOHQILTATEATHHTLNN 537
QY      182 YQYRVARATIRYPLVNAVGGVAISIFWPQTTTPTSDVMSITSDVRLVQPGIA 241
Db      538 AAGQLHVNGTLDLTAQHLNNAAGHLLTTG--PQSATL-TIADLLDNTSGTL-----ASA 588
QY      242 SELVIPERLHYRNGQWRSVETSGVAEEBATSGLVMLCIHGLSVNSYNTPTYGALGLLD 301

```

Db 589 GSLTLTAAPL-----DITDGTQSGGQGPLHMDAATLTAHKGTLTSD 630
QY 302 FALELEFRNLTPGNTNTRVRSYSSTAR-----HRLRGA----- 335
Db 631 -----TLTLGTHDLSHATTTAAQIILHTGDLTTAGGHLTAYGTHTLQLEARTLD 682
QY 336 -----DGTAELETTAATRFMKDLYFTSTNGVGVIGRGIALTLNADTLTGGPLTEL 387
Db 683 NTGGTIATNGTLDLHTAALDNTGTLHSTAT---GNRLDITDTLTNTAGHLLNGPTTL 739
QY 388 I---SSAGGOLFYSRPVV-----SANGPTVKLYTSVENAQDQKGAIPH 429
Db 740 TTGTWNTTNGHLQITGPATLHATLTLNDRGILHTATGDLRLITGLINN--QNGLLTTT 797
QY 430 D-IDLGESRVVIQDYNQH-EQDRPTSPAP-----SRPFSVLRANDVWL----- 473
Db 798 DALTLTAASLL-----NQGTLDAAAGPAHLTLTGLDNTAGLLQTAHLTLWLTSAGLTNR 852
QY 474 -SLTAAEYDQSTYSS-----TGPV-----YVSDSVTL 500
Db 853 GTLTASQLTLDTCGHTLDNTRGRGLTTGNTGLHSGPLDNTAGLLQTAALTTDTGAATL 912
QY 501 VNVATGAQAVARSIDMTKVLDCRPLSTIQYQSKTFVFLRGLKLSWEAGTTKAGYPYN 560
Db 913 TNRDGGALLAATLDTTATLNRG-GTIDSQTAT-----HLHTTLDNTAGHISS 963
QY 561 YNTTASDQLLVENAAAG 576
Db 964 SGTLQIDGTLTWTGG 979

RESULT 10

B82519
hemagglutinin-like secreted protein XF2775 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: B82519

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-3455 <SIM>

A;Cross-references: GB:AE004082; GB:AE003849; NID:G9108003; PIDN:AAF85560.1; GSPDB:GN001

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorzy, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.B.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv

M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A;Reference number: A59328

A;Contents: annotation

C;Genetics:

A;Gene: XF2775

Query Match 4.0%; Score 135.5; DB 2; Length 3455;

Best Local Similarity 21.9%; Pred. No. 8;

Matches 135; Conservative 64; Mismatches 216; Indels 201; Gaps 27;

QY 97 ASRRPPTTAGAAPLTAAPAHPTP-----PVP-----DVDSRGAIL 132

Db 429 ASPADPT-----FVASVPATDTPPTGTPPIPSHPQAPVLAAGRLTLENDIDNRGGHI 483

QY 133 RQYNLSTSPLTSSVATGNLVLYAAPLSPL-LPQDGTNTH-----IMATEA-----SN 181
Db 484 TAGGADAILTTLDNRDGT-----AALNRLTQGLQDQHGILLTATEATHHTLNN 537
QY 182 YQYRVARATIRYRPLVFNAGVYASISFWPQTPTTPTSDVMSNSTSTSDRILLVOPGIA 241
Db 538 AAGQLHVNGTLDLTAQHLNNSAGHLITG--FQSATL-TIADLLDNTSGTL-----ASA 588
QY 242 SELVTPSERLHYRNQWRSVETSGVABEATSGLYMLCTHGSIVNSYNTPTTGALGLLD 301
Db 589 GSLTLTAAPL-----DITDGTQSGGQGPLHMDAATLTAHKGTLTSD 630
QY 302 FALELEFRNLTPGNTNTRVRSYSSTAR-----HRLRGA----- 335
Db 631 -----TLTLGTHDLSHATTTAAQIILHTGDLTTAGGHLTAYGTHTLQLEARTLD 682
QY 336 -----DGTAELETTAATRFMKDLYFTSTNGVGVIGRGIALTLNADTLTGGPLTEL 387
Db 683 NTGGTIATNGTLDLHTAALDNTGTLHSTAT---GNRLDITDTLTNTAGHLLNGPTTL 739
QY 388 I---SSAGGOLFYSRPVV-----SANGPTVKLYTSVENAQDQKGAIPH 429
Db 740 TTGTWNTTNGHLQITGPATLHATLTLNDRGILHTATGDLRLITGLINN--QNGLLTTT 797
QY 430 D-IDLGESRVVIQDYNQH-EQDRPTSPAP-----SRPFSVLRANDVWL----- 473
Db 798 DALTLTAASLL-----NQGTLDAAAGPAHLTLTGLDNTAGLLQTAHLTLWLTSAGLTNR 852
QY 474 -SLTAAEYDQSTYSS-----TGPV-----YVSDSVTL 500
Db 853 GTLTASQLTLDTCGHTLDNTRGRGLTTGNTGLHSGPLDNTAGLLQTAALTTDTGAATL 912
QY 501 VNVATGAQAVARSIDMTKVLDCRPLSTIQYQSKTFVFLRGLKLSWEAGTTKAGYPYN 560
Db 913 TNRDGGALLAATLDTTATLNRG-GTIDSQTAT-----HLHTTLDNTAGHISS 963
QY 561 YNTTASDQLLVENAAAG 576
Db 964 SGTLQIDGTLTWTGG 979

RESULT 11

T21460

hypothetical protein ZK945.10 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T21460

R;Wilkinson, J.

submitted to the EMBL Data Library, March 1995

A;Reference number: Z19425

A;Accession: T21460

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-796 <WTL>

A;Cross-references: EMBL:Z48582; PIDN:CAA88469.1; GSPDB:GN00020; CESP:ZK945.10

A;Experimental source: clone F27E5

C;Genetics:

A;Gene: CESP:ZK945.10

A;Map position: 2

A;Introns: 62/3; 124/3; 169/1; 216/1; 262/3; 371/1; 581/3; 608/1

Query Match 3.9%; Score 133; DB 2; Length 796;

Best Local Similarity 20.6%; Pred. No. 1.4;

Matches 121; Conservative 61; Mismatches 228; Indels 176; Gaps 24;

QY 63 APDVTAAGAGPRVQPARPLGSAMRQOAPAVASRRRPTTAGAAPLTAAPAHPTPPV 122

Db 242 AKQFSMRTSQSPTLRRMKRDAGDNTCDYTIESTSTSTTPTTT-----TVTSTVTSTTV 296

QY 123 PDVDSRGAILLRQYNLSTSPLTSSVATGNLVLYAAPLPLQDGTNTHIMATEAGNY 182

Db 297 P-----TSTSTVTWTAMSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 338


```
QY 109 PUTAVAPAHDTPPVPVDVGRGAILRR-----QYNLSTSP-----TSS 146
Db 28 PTTSSVSSADPTTSADPTTSVQTTSAGPSNNIGNSTLANSTTFAVSTSDIDPTSSSD 87
QY 147 VATGTNLVLVYAAPLSPLLPQD---CTNTHIMATEASNYAQYVARATIRYRPLVNAV 203
Db 88 VIITSTVQTTSIEPTTSLVSSNDITTSNLINISVVIISTDSTSLIESTTTVGHASSVA 147
QY 204 G-YAISISFPWPTTTPTSDMNSITSDVRLIVQPGIAS---ELVIPSERLHYRQGW 259
Db 148 GMYRTSSDEVITSTOPTS-----SSDVATSDPTSSSAVTTLVDPPTSV----- 192
QY 260 SVETSGVAEEATSGVLMLCIHGSLVNSYNTPTGTALGLLDPALEFRNLTPGNVTR 319
Db 193 -VISTSDVOTSSSDVATSDPTTSVISTSSADPTTSA-----DSTTSTVQTTS 239
QY 320 VSRYSSTARHLRRGADGTAEELTTAATRFMKDLYFTSTNGVGEIGRGIALTLFNLADTL 379
Db 240 VDTSSVSSAPVDPASSVSLTSSVPTSSSTVITISANGSGATLA-----AQIT 289
QY 380 LGLPLTELSSAGGOLFGRP-----YVSANGPTVKLYTSVE-NAQDQKGIAPHD 430
Db 290 SIDPVSSIVSSGATTIISASIDPASSVVSSTSSPTSFIVSSTSVSTRPSGPTTSTD 349
QY 431 IDIGESRVVIQDYDQHEQDRPTSPAPRPFSVLKRVANDVLMLSLTAAYDQSTYGSSTG 490
Db 350 LATFSDTIILRVTTSTSQDTQTVSSSLTDMVSSSTGSDLSVSSIQRSQVDPSTFAVNS 409
QY 491 PVVVSQVTLVNVATGAQAVARSLDMTK-----VTLDGRPLSTIQVSKTFFV 538
Db 410 PVYPTASTG--SVSTGTPASELSLSRQOGISATSSSVITLTPVDSASSSSSSATSII 467
QY 539 LP-----LRGKLSFWEA-GTTKAGYP-----YNYNTASDQLIV-ENAAHRVAIS 582
Db 468 KPNWPSSNDKQTSQSVVVDAFQSTKSSYPSTISADPTTLASENGLVGSSSAHPITLD 527
QY 583 -TYTTSLGAGPVSISAVAVLAPHSA 607
Db 528 RTYASAHASVTDIVSRVTDSTRHTTL 553

RESULT 14
S59310
probable membrane protein YMR317w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YW924.09
C;Species: Saccharomyces cerevisiae
C;Date: 29-Nov-1995 #sequence_revision 23-Feb-1996 #text_change 19-Apr-2002
C;Accession: S59310
R;Churcher, C.M.
submitted to the EMBL Data Library, September 1995
A;Reference number: S59302
A;Accession: S59310
A;Molecule type: DNA
A;Residues: 1-1104 <CHU>
A;Cross-references: EMBL:Z54141; GSPDB:GNO0013; MIPS:YMR317w
A;Experimental source: strain AB972
C;Genetics:
A;Gene: MIPS:YMR317w
A;Cross-references: SGD:S0004936
A;Map position: 13R

Query Match 3.8%; Score 130.5; DB 2; Length 1104;
Best Local Similarity 18.7%; Pred. No. 3.2; Indels 121; Gaps 22;
Matches 113; Conservative 109; Mismatches 262;

QY 63 APDVTAAGAGPRVRQPARPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPA-----H 117
Db 148 SPQTSSSNGG-----SSSEPLGK-----SSVLETTASSDITVATSTFTTLT 191
QY 118 DTPPPVDVDSRGAIL-----RQYNI.LST-----SPLTSSVA-----TGT 151
Db 192 DVSSSPKISSSGSAVTSVGTTSDAKSEVFSSSTSDVSSLLSSTSSPASSTISETLPFSST 251
```

```
QY 152 NLVLVYAAPLSPLLPQDGTNTHIMATEASNYAQYVARATIRYRPLVNAVGGYAIISI-- 209
Db 252 ILUSTSSPVSSSEAP--SATSSSVSEASSTSSSVSEAPLATSSVVSSEAPSTSSVVS 309
QY 210 SEWPQTTTTPTSDMNSITSDVRLIVQPGIASLSEVIPSE-----RLH 252
Db 310 SEAPSTSSSVSEISSTTSSVS--SEAPLATSSVVSSEAPSTSSSVSSEAPSTSSS 367
QY 253 YNQGWRSVETSGVAEEATSGVLMLCIHGSLVNSYNTPTGTALGLLDPALEFRNL 312
Db 368 VSSEAPSTSSSVSEISSTKSSVMSSEVSSATSSSLVSEAPSAISL--ASSRLPSSKN 425
QY 313 PGNTNTRVRSYSTARHLRRGADGTAE--LTTAATRFMKDLYFTSTNGVGEIGRGIA 369
Db 426 TSVTSLVATEASSVTSRLRPSSETLASNIIESSLTGVNSTVSTTSAASSTLGSKVS 485
QY 370 LTLFNLADTLGLGLPTELISAGGOLFYSRPPVVSANGPTVKLYTSVENNAQDQKGIAP- 428
Db 486 SSNSRMATKSTSTSSDL--SKSSVIFGNSSTVTTSPSASISLTAS-----PLPS 533
QY 429 --HDIDLGESRVVIQDYDQHEQDRPTSPAPRPFSVLKRVANDVLMLSLTAAYDOSTYG 486
Db 534 VMSDITSSSEASSISNL-----ASSAPSDDNNSTIAS-----ASLLVTKIKSVWS 579
QY 487 SSTGPVVVSDSVTLNVNATGA-----QAVARSLDMTKVT-LDGRPLSTIQVSK-TFFV 538
Db 580 SVVSSITSETNESNLATSSSTLLSNKATARSLSNATASNVPTGTPTSSWSHSTSVI 639
QY 539 LPARGKLSFWEAGTTKAGYPYNYNTASDQLIVENAAHRVAISYTTSLGAGPVSISAV 598
Db 640 TP-----GFSTSGASLAIN-STVVSSSLAGYSPSTPSSPTTSTLTSEAPSTVSSM 690
QY 599 AVLAP 603
Db 691 TTSAP 695

RESULT 15
S57180
probable membrane protein YJR151c - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein J2223; serine/threonine-rich protein YJR151c
C;Species: Saccharomyces cerevisiae
C;Date: 23-Aug-1995 #sequence_revision 08-Sep-1995 #text_change 19-Apr-2002
C;Accession: S57180
R;Scarcez, T.
submitted to the Protein Sequence Database, September 1995
A;Reference number: S57169
A;Accession: S57180
A;Molecule type: DNA
A;Residues: 1-1161 <SCA>
A;Cross-references: EMBL:Z49651; NID:g1015902; PID:g1015903; GSPDB:GNO0010; MIPS:YJR151c
C;Genetics:
A;Gene: SGD:DAN4; MIPS:YJR151c
A;Cross-references: SGD:S0003912
A;Map position: 10R
C;Keywords: transmembrane protein

Query Match 3.8%; Score 130.5; DB 2; Length 1161;
Best Local Similarity 17.2%; Pred. No. 3.4; Indels 133; Gaps 21;
Matches 104; Conservative 95; Mismatches 274;

QY 57 HPTNPFAPDVTAAAG-----AGPRVRQPARPLGSAWRDQAPAVASRRR--- 101
Db 59 HKETETYSSEIAAAVFDYGDFTTTLTGISGDEVTRMI--TGVPMYSTRKLPKPAISSALSKDG 116
QY 102 -----PTTAGAAPLTAVAPAHDTPPVDVSRGAILRQYNI.LSTSSVA-----TGTN 152
Db 117 IYTAIPTSTSTTTTKSSTSTPTTITSTTSTTSTTSTTSTTSTTSTTSTTSTTSTTSTT 176
QY 153 LVLYAAPLSPLLPQDGTNTHIMATEASNYAQYVARATIRYRPLVNAVGGYAIISFW 212
Db 177 STPTTSTTSTTPTTSTTST-----TPTTSTTSTTPTTSTTSTTSTTSTT-----TTS 224
```


Blank Sheet

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	3408	100.0	660	1	VST2_HSVBU	P29326 hepatitis e
2	3391	99.5	660	1	VST2_HSVBU	P33426 hepatitis e
3	3367	98.8	660	1	VST2_HEWMY	Q04611 hepatitis e
4	3194.5	93.7	659	1	VST2_HEWME	Q03500 hepatitis e
5	2446	71.8	485	1	VST2_HEVRH	Q02070 hepatitis e
6	143	4.2	1041	1	EGT2_YEAST	P42835 saccharomyc
7	137.5	4.0	1083	1	T2D3_HUMAN	Q00268 homo sapien
8	137	4.0	5703	1	MUSE_HUMAN	Q9bc84 homo sapien
9	133	3.9	670	1	YFGG_SCHPO	O13854 schizosacch
10	133	3.9	1802	1	HKRL_YEAST	P41809 saccharomyc
11	133	3.9	3178	1	YS89_CABEL	Q09624 caenorhabdi
12	131	3.8	1322	1	YAG3_YEAST	P39712 saccharomyc
13	130.5	3.8	768	1	AMY1_SACDI	P29760 saccharomyc
14	130.5	3.8	1161	1	DAN4_YEAST	P47179 saccharomyc
15	129.5	3.8	1321	1	PGCN_HUMAN	O14594 homo sapien
16	129.5	3.8	1367	1	AMVH_YEAST	Q08640 saccharomyc
17	129	3.8	1015	1	PTPX_HUMAN	Q92932 homo sapien
18	128.5	3.8	1537	1	FLOI_YEAST	P32768 saccharomyc
19	128	3.8	767	1	AMVH_SACDI	P04065 saccharomyc
20	127.5	3.7	518	1	VL2_HPVA7	P22425 human papil
21	126.5	3.7	351	1	COAT_TCV	P06663 turnip crin
22	126.5	3.7	1306	1	MSB2_YEAST	P32334 saccharomyc
23	126	3.7	3664	1	MINT_HUMAN	Q95t58 homo sapien
24	125.5	3.7	725	1	AGAL_YEAST	P32323 saccharomyc
25	125	3.7	697	1	BYN_DROME	P55965 drosophila
26	125	3.7	1054	1	S24A_ARATH	Q9sfu0 arabidopsis
27	123.5	3.6	797	1	VGLX_HVEVB	P28968 equine herp
28	123.5	3.6	1119	1	ALS3_CANAL	O74623 candida alb
29	123.5	3.6	1140	1	YV96_YEAST	Q04893 saccharomyc
30	123	3.6	2030	1	HFCL_MESAU	P54611 mesocricetu
31	122	3.6	472	1	SK14_DROME	P40656 drosophila
32	122	3.6	3354	1	CADN_HUMAN	Q9H251 homo sapien
33	121.5	3.6	7743	1	POZ1_HUMAN	P14859 homo sapien

```

Db 121 PVPDVSRAILRRQYNLSTPLTSSVATGTLNVLVYAAPLSPLLPLQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRLPVNAVGGYAISSFPQTPTTPTSDVDMNSITSDVRILVQPGI 240
Db 181 NYAQYRVARATIRYRLPVNAVGGYAISSFPQTPTTPTSDVDMNSITSDVRILVQPGI 240
QY 241 ASELVIPSERLHYRNQGRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGTALGLL 300
Db 241 ASELVIPSERLHYRNQGRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGTALGLL 300
QY 301 DFALFELFNLTGNTNTRVSRYSSTARHLRRGADGTAEELTTTAATRFMKDLYFTSTNG 360
Db 301 DFALFELFNLTGNTNTRVSRYSSTARHLRRGADGTAEELTTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTELISAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEODRPTSPAPSRPFVLRANDVWLJSLTAAEY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEODRPTSPAPSRPFVLRANDVWLJSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTIQQYSKTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTIQQYSKTFFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHGRAHVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHGRAHVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDLDYPARAHTPDDFCPCRPGLGQCAFOSTVAELQRLKMKVGKTREL 660
Db 601 LAPHSALELLEDLDYPARAHTPDDFCPCRPGLGQCAFOSTVAELQRLKMKVGKTREL 660

RESULT 2
VST2_HEVPA
ID VST2_HEVPA STANDARD; PRT; 660 AA.
AC P33426;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=33774;
RN [1]
RP MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M80581; AAA45727.1; ..
CC InterPro: IPR004261; SP2.
CC DR InterPro: IPR008975; Viral_cap_coat.
CC DR Pfam: PF03014; SP2; 1.
CC SSignal.
CC FT SIGNAL
CC CHAIN 23 660 BY SIMILARITY.

```

```

SQ SEQUENCE 660 AA; 70980 MW; 8085BC53CPB4FD3 CRC64;
Query Match 99.5%; Score 3391; DB 1; Length 660;
Best Local Similarity 99.4%; Pred. No. 2,7e-207;
Matches 656; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 MRPRPILLLLMLPMLPAPPQPSGRRRGRSGSGGFGMDRVDSQFALPYIHTPN 60
Db 1 MRPRPILLLLMLPMLPAPPQPSGRRRGRSGSGGFGMDRVDSQFALPYIHTPN 60
QY 61 PFAPDVTAAGAGPRVRQPARPLGSARWQAQAPAVASRRRPTTAGAAPTAVAPADTP 120
Db 61 PFAPDVTAAGAGPRVRQPARPLGSARWQAQAPAVASRRRPTTAGAAPTAVAPADTP 120
QY 121 PVPDVSRAILRRQYNLSTPLTSSVATGTLNVLVYAAPLSPLLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRAILRRQYNLSTPLTSSVATGTLNVLVYAAPLSPLLPLQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRLPVNAVGGYAISSFPQTPTTPTSDVDMNSITSDVRILVQPGI 240
Db 181 NYAQYRVARATIRYRLPVNAVGGYAISSFPQTPTTPTSDVDMNSITSDVRILVQPGI 240
QY 241 ASELVIPSERLHYRNQGRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGTALGLL 300
Db 241 ASELVIPSERLHYRNQGRSVETSGVAEEATSGLVMLCIHGSPVNSYNTPTGTALGLL 300
QY 301 DFALFELFNLTGNTNTRVSRYSSTARHLRRGADGTAEELTTTAATRFMKDLYFTSTNG 360
Db 301 DFALFELFNLTGNTNTRVSRYSSTARHLRRGADGTAEELTTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTELISAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEODRPTSPAPSRPFVLRANDVWLJSLTAAEY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEODRPTSPAPSRPFVLRANDVWLJSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTIQQYSKTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTIQQYSKTFFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHGRAHVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHGRAHVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDLDYPARAHTPDDFCPCRPGLGQCAFOSTVAELQRLKMKVGKTREL 660
Db 601 LAPHSALELLEDLDYPARAHTPDDFCPCRPGLGQCAFOSTVAELQRLKMKVGKTREL 660

RESULT 3
VST2_HEVMY
ID VST2_HEVMY STANDARD; PRT; 660 AA.
AC Q04611;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227573; PubMed=8470371;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikihisa T., Winn K.;
RT "Sequence and gene structure of the hepatitis E virus isolated from
RT Myanmar.";
RL Virus Genes 7:95-109(1993).
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC

```

CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D10330; BAA01174.1; -;
DR InterPro: IPR004261; SP2;
DR Pfam: PF03014; SP2; 1.
KW SIGNAL.
FT CHAIN 1 22 BY SIMILARITY.
FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70998 MW; 3A82A4EA25C6253 CRC64;

Query Match 98.8%; Score 3367; DB 1; Length 660;
Best Local Similarity 98.9%; Pred. No. 8.8e-206;
Matches 653; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRPRDILLILLMLPMLPAPPQSGRRRRSGSGGFGWDRVDSQPPFAIPYIHPTN 60
DB 1 MRPRDILLILLMLPMLPAPPQSGRRRRSGSGGFGWDRVDSQPPFAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHPTP 120
DB 61 PFAPDVTAAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHPTP 120
QY 121 PVPDVSRRGAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLLLPLQDGTNTHIMATEAS 180
DB 121 PVPDVSRAAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLLLPLQDGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYPLVPNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIIVQPGI 240
DB 181 NYAQVRVARATIRYPLVPNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIIVQPGI 240
QY 241 ASELVIPSERLHYRNQGWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGALGLL 300
DB 241 ASELVIPSERLHYRNQGWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGALGLL 300
QY 301 DFALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
DB 301 DFALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPLELISAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNLADTLGLLPLELISAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSPFVSRLRANDVWLSLTAAEY 480
DB 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSPFVSRLRANDVWLSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGDRPLSTIQYKTFVFLP 540
DB 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGDRPLSTIQYKTFVFLP 540
QY 541 LRKLSFWEAGTTKAGYPYNTNTASDQLLVENAGHRAVAISTYTTSLGAGPVSISAVV 600
DB 541 LRKLSFWEAGTTKAGYPYNTNTASDQLLVENAGHRAVAISTYTTSLGAGPVSISAVV 600
QY 601 LAPHSALEEDTLDPACAHTEFDDFCPCPLGIGCCAFOSTVAELORLKKVCKTREL 660
DB 601 LAPHSALEEDTLDPACAHTEFDDFCPCPLGIGCCAFOSTVAELORLKKVCKTREL 660

RESULT 4
VST2 HEVME
ID VST2 HEVME STANDARD; PRT; 659 AA.
AC Q03500;
DC 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor.
OS Hepatitis E virus (strain Mexico) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31768;
RN [1] SEQUENCE FROM N.A.
RX MEDLINE=93079857; PubMed=1448913;
RA Huang C.C., Nguyen D., Fernandez J., Yun K.Y., Fry K.E.,
RA Bradley D.W., Tam A.W., Reyes G.R.;
RT "Molecular cloning and sequencing of the Mexico isolate of hepatitis
RT E virus (HEV).";
RL Virology 191:550-558 (1992).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M74506; AAA45732.1; -;
DR PIR: B44212; B44212.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
KW SIGNAL.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 659 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 659 AA; 70640 MW; CF75E75EFD8FBE2C CRC64;

Query Match 93.7%; Score 3194.5; DB 1; Length 659;
Best Local Similarity 93.0%; Pred. No. 7.3e-195;
Matches 614; Conservative 21; Mismatches 24; Indels 1; Gaps 1;

QY 1 MRPRILLILLMLPMLPAPPQSGRRRRSGSGGFGWDRVDSQPPFAIPYIHPTN 60
DB 1 MRPRILLILLMLPMLPAPPQSGRRRRSGSGGFGWDRVDSQPPFAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHPTP 120
DB 61 PFAPDVTAAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHPTP 120
QY 121 PVPDVSRRGAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLLLPLQDGTNTHIMATEAS 180
DB 121 PVPDVSRRGAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLLLPLQDGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYPLVPNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIIVQPGI 240
DB 181 NYAQVRVARATIRYPLVPNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIIVQPGI 240
QY 241 ASELVIPSERLHYRNQGWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGALGLL 300
DB 241 ASELVIPSERLHYRNQGWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGALGLL 300
QY 301 DFALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
DB 301 DFALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPLELISAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNLADTLGLLPLELISAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSPFVSRLRANDVWLSLTAAEY 480
DB 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSPFVSRLRANDVWLSLTAAEY 480

QY 481 DOSTYSSGPVYVSDSVTLNVNATGAQAVARS�DWTKVTLDGRPLSTIQOYKSTFFVL 540
DB 480 DOSTYSSGPVYVSDSVTLNVNATGAQAVARS�DWSKVTLDGRPLPTVEQYKSTFFVL 539
QY 541 LRKLSFWAGTTKAGYPYNYNTTASDQLLVNNAAGHRVAISTYTTSLGAGPVSISAVAV 600
DB 540 LRKLSFWAGTTKAGYPYNYNTTASDQLLVNNAAGHRVAISTYTTSLGAGPVSISAAV 599
QY 601 LAPHSALLLEDLTDPARAHTFDDFCPCRLGLOGCAFSQSTVABLRQKMKVKGKREL 660
DB 600 LAPRSALLLEDTFDYPGRAHTFDDFCPCRALGLOGCAFSQSTVABLRQKMKVKGKREL 659
RESULT 5
VST2 HEVRH STANDARD; PRT; 485 AA.
AC Q00270;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE Structural protein 2 (Fragment).
OS Hepatitis E virus (isolate Rhesus) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31766;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K.,
RA Win K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90274; BAA20910.1; --
CC InterPro; IPR004261; SP2.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF03014; SP2; 1.
CC NON_TER 1 485
CC NON_TER 485 485
CC SEQUENCE 485 AA; 52317 MW; 5A4FD9273AC74F94 CRC64;
Query Match 71.8%; Score 2446; DB 1; Length 485;
Best Local Similarity 99.0%; Pred. No. 1.2e-147;
Matches 480; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 133 RQYNLSTSPSTSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEASNAQYVARATI 192
DB 1 RQYNLSTSPSTSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEASNAQYVARATI 60
QY 193 RYRLVNAVGGYIAISISFWPQTPTTPTSDVMSNITSDVRLVQPGIASLVIPSERLH 252
DB 61 RYRLVNAVGGYIAISISFWPQTPTTPTSDVMSNITSDVRLVQPGIASLVIPSERLH 120
QY 253 YRNOGWSRVETSGVAEBEATSGLVMLCIHGSLVNSYNTPTGTGALLDFALEFRNL 312
DB 121 YRNOGWSRVETSGVAEBEATSGLVMLCIHGSPVNSYNTPTGTGALLDFALEFRNL 180
QY 313 PGNTNTRSVRSYSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNGVBIGRGIALTL 372
DB 181 PGNTNTRSVRSYSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNGVBIGRGIALTL 240
QY 373 FNLADTLGLLPTELISAGGQLFYSRFPVWSANGEFTVKLYTSVENAQDQKGIAPHDID 432
DB 241 FNLADTLGLLPTELISAGGQLFYSRFPVWSAHGFTVKLYTSVENAQDQKGIAPHDID 300

QY 433 LGSRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVLWLSITAAEYDQSTVGSSTGPV 492
DB 301 LGSRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVLWLSITAAEYDQSTVGSSTAPV 360
QY 493 YVSDSVTLNVNATGAQAVARS�DWTKVTLDGRPLSTIQOYKSTFFVLPLRGKLSFWEAGT 552
DB 361 YVSDSVTLNVNATGAQAVARS�DWTKVTLDGRPLSTIQOYKSTFFVLPLRGKLSFWEAGT 420
QY 553 TKAGYPYNYNTTASDQLLVNNAAGHRVAISTYTTSLGAGPVSISAVAVLAPHSAALLED 612
DB 421 TKAGYPYNYNTTASDQLLVNNAAGHRVAISTYTTSLGAGPVSISAVAVLAPHSAALLED 480
QY 613 TLDYP 617
DB 481 TLDYP 485
RESULT 6
EGT2_YEAST STANDARD; PRT; 1041 AA.
ID EGT2_YEAST
AC P42835;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE EGT2 protein precursor (Early G1 transcript 2).
GN EGT2 OR YNL327W OR N0320.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1676;
RX MEDLINE=95373280; PubMed=7645347;
RA Maftahi M., Nicaud J.-M., Levesque H., Gaillardin C.;
RT "Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV
RT identifies the RPD3, PAS8 and KRE1 loci, five new open reading
RT frames.";
RL Yeast 11:567-572(1995).
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=96251274; PubMed=8668141;
RA Kovacech B., Nasmith K., Schuster T.;
RT "EGT2 gene transcription is induced predominantly by Swi5 in early
RT G1.";
RT Mol. Cell. Biol. 16:3264-3274(1996).
CC -1- FUNCTION: Seems to be involved in the correct timing of cell
CC separation after cytokinesis, as separation of mutant daughter
CC cells is delayed. Could either be an enzyme necessary for glucans-
CC degradation of the cell wall at the neck region between mother and
CC daughter cells or a regulatory protein controlling this metabolic
CC step. Exclusively expressed between the end of mitosis and early
CC G1; inactivated before cells pass start.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z45259; CAA86371.1; --
CC EMBL; Z71603; CAA96259.1; --
CC PIR; S55862; S55862.
CC Germonline; 143333; --
CC SGD; S0005271; EGT2.
CC GO; GO:0009277; C:cell wall (sensu Fungi); IDA.
CC GO; GO:0000910; P:cytokinesis; IMP.
CC Glycoprotein; Repeat; Signal; Cell cycle; Cell division; Mitosis.
FT SIGNAL 1 20
FT CHAIN 21 1041 EGT2 PROTEIN.

FT	DOMAIN	200	POLY-SER.	203	
FT	DOMAIN	381	POLY-SER.	384	
FT	DOMAIN	388	POLY-SER.	395	
FT	DOMAIN	490	POLY-SER.	493	
FT	DOMAIN	586	POLY-SER.	589	
FT	REPEAT	457	1-1.	492	
FT	REPEAT	577	1-2.	606	
FT	REPEAT	613	1-3.	647	
FT	REPEAT	716	1-4.	745	
FT	REPEAT	773	1-5.	802	
FT	REPEAT	811	1-6.	840	
FT	REPEAT	849	1-7.	886	
FT	REPEAT	887	1-8.	924	
FT	REPEAT	925	1-9.	962	
FT	CARBOHYD	65	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	103	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	161	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	175	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	249	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	332	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	401	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	435	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	465	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	485	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	506	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	526	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	544	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	556	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	635	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	636	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	657	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	709	N-LINKED	(GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	756	N-LINKED	(GLCNAC. .)	(POTENTIAL).
SQ	SEQUENCE	1041 AA, 108494 MM, 01FEFCF8EA8744CD CRC64;			
Query Match 4.2%; Score 143; DB 1; Length 1041;					
Best Local Similarity 20.4%; Pred. No. 0.21;					
Matches 144; Conservative 94; Mismatches 238; Indels 230; Gaps 31;					
QY	48	SQFFAIPYHPNPAPDVTAAAGAGPRVQPARPLGSAWRQAPQAPAVASRRPTT---	104		
Db	340	SSQSAIPEV-----AVTASLSSG-----ILSSTVDGASTADASMSAVSTVSS	382		
QY	105	---AGAPLTAAPAHDPDPVDSRGAILLRQVNLSTSLTSSVATGTLNLYAAPL	160		
Db	383	SSQASSSSISLAFSSSN-----STFTTSSLSASATETYSIISASGI	425		
QY	161	SPLLPLQDGTNTHIMATEASNAQYRVAR---ATIRYRLVNPVNAVGGYAISSFQPTTT	217		
Db	426	SVTQASYIDNSTTAVTQSTSTIAVSAAEKLSSLTLY-----TSNVTIVSSATQHTT	478		
QY	218	TPTSVDMSNITGDVRLVQPIASBELVTPSERLHYRNOGWRSVETSGVABEATSGLVN	277		
Db	479	TPSYYS-NSTTLSSSVL-----ESVSSPYLIANTTVSG---	511		
QY	278	LCIHGSLVNSYNTPY-----TGALGLDFALEFRNLTPGNTNTRVRSYSSTAR	328		
Db	512	-----ASSASQSTNPPVNSNTSSATQATATAPFAINITGTSISSITNI-----SSVSSST-	564		
QY	329	HLRRG-----ADGTAELTTTAAATRFMKL-----YFTSTNGVGEIQRG---	367		
Db	565	SSLSSGPFVNSNATAVAGSVILITTTESAQLTEIGSLIPTITTTSTTSGTDKTSNKVA	624		
QY	368	-----TALTFLNADTLGLPLTELSSA-----GGQL-----PYSRPVSNAGE	407		
Db	625	SSTEIAQSVNNSLSVSTINTNAATAAANARNATFTTHATSGSLQPSYHSSLSLSTID	684		
QY	408	PTVKLYTSVENAQDQKGIAPHDIDIGESRVVIQDYDQNEHQDRPTSPAPRPSVLRA	467		
Db	685	---TKVTATTSTSRGSSSLAFTTGLNQSVVTGTD-----KSDTYSVTS	727		
QY	468	NDVLMLSLTAAYDQ-----STY-----GSSTGPVVVSDSVTLNVNATG-----AQAVAR	512		

Db	728	TE-----SAQVTEYDSLLPISTLTKPTVVTCTSRNSTFSMVSSTKLTATATDKGDAYSVIS	783		
QY	513	SLDWTKYVLDR--PLSTIQ-----QYKTFVLPGRK	544		
Db	784	STQSAQVTEYGSMLPISLETPTVMSTDESGYFTLTCTESGOATEYGLSIPSTLDGS	843		
QY	545	LSFWEAG-----TTKAGYPYNYNTTASDOLLVANAAGHVAIST-----YTTSLG	589		
Db	844	VYITFTGESVVVGYSTTVGAQAQHTSL---VPVSTIKGSKTSLSTESVAGYSTTVG	900		
QY	590	AG-----PVSI---SAVAVLAPHASALALLEDTLDYPARA-HT	622		
Db	901	AAQAQHTSLVPVSTIKGSKTSLSTESVAGYSTTVDSQAQAEHT	946		
RESULT 7					
T2D3 HUMAN					
ID	T2D3 HUMAN	STANDARD;	PRT; 1083 AA.		
AC	O00268; Q99721; Q9BR40; Q9BX42;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	Transcription initiation factor TFIID 135 kDa subunit (TAFII-135) (TAFII135) (TAFII-130) (TAFII130).				
GN	TAF4 OR TAF4 OR TAF2C1 OR TAF2C1 OR TAFII135 OR TAFII130.				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
OX	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=973336072; PubMed=9192867;				
RA	Mengus G., May M., Carre L., Chambon P., Davidson I.;				
RT	"Human TAF(II)135 potentiates transcriptional activation by the AP-2s				
RT	of the retinoic acid, vitamin D3, and thyroid hormone receptors in				
RT	mammalian cells.";				
RL	Genes Dev. 11:1381-1395(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21638749; PubMed=11780052;				
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,				
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,				
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,				
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,				
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,				
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,				
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,				
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,				
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,				
RA	Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,				
RA	Hackle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,				
RA	Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,				
RA	Lehvaeslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,				
RA	Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,				
RA	Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,				
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,				
RA	Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,				
RA	Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,				
RA	Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,				
RA	Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,				
RA	Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,				
RA	Whithead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,				
RA	Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,				
RA	Rogers J.;				
RT	"The DNA sequence and comparative analysis of human chromosome 20.";				
RL	Nature 414:865-871(2001).				
RN	[3]				
RP	SEQUENCE OF 105-1083 FROM N.A., AND PARTIAL SEQUENCE.				
RX	MEDLINE=97098442; PubMed=8942982;				
RA	Tanese N., Saluja D., Vassallo M.F., Chen J.-L., Admon A.;				
RT	"Molecular cloning and analysis of two subunits of the human TFIID				
RT	complex: hTAFII130 and hTAFII100.";				


```
FT CARBOHYD 4965 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 4987 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5037 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5052 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5156 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5427 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5467 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5506 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5507 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5543 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5553 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5604 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5618 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5662 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT VARIANT 5137 T -> S (in dbSNP:2672788).
                          /FTid=VAR_014123.
FT CONFLICT 34 G -> E (IN REF. 2).
FT CONFLICT 95 FPGLCN -> LPCLCK (IN REF. 2).
FT CONFLICT 104 S -> C (IN REF. 2).
FT CONFLICT 142 E -> K (IN REF. 1).
FT CONFLICT 225 R -> S (IN REF. 2).
FT CONFLICT 330 PL -> T (IN REF. 2).
FT CONFLICT 337 E -> N (IN REF. 2).
FT CONFLICT 356 E -> K (IN REF. 2).
FT CONFLICT 362 G -> R (IN REF. 2).
FT CONFLICT 369 MISSING (IN REF. 2 AND 3).
FT CONFLICT 374 D -> N (IN REF. 2).
FT CONFLICT 393 RT -> TR (IN REF. 2).
FT CONFLICT 468 RK -> GR (IN REF. 2).
FT CONFLICT 512 L -> P (IN REF. 2).
FT CONFLICT 585 GAA -> AH (IN REF. 3).
FT CONFLICT 601 A -> S (IN REF. 3).
FT CONFLICT 628 DP -> RS (IN REF. 2).
FT CONFLICT 633 F -> L (IN REF. 2).
FT CONFLICT 676 A -> P (IN REF. 3).

Query Match
Best Local Similarity 4.0%; Score 137; DB 1; Length 5703;
Matches 143; Conservative 20.1%; Pred. No. 5.1;
Matches 143; Conservative 72; Mismatches 257; Indels 240; Gaps 33;

QY 52 AIPYIH--PTNFPADVTAAGAPRVQPARPLGSAWRDQAPAVASRRRTTAGAAP 109
DB 3192 ALPALRSTATTPTATSTVATL-----PSSSLGTAW-----TRLSQTTTPTAT 3232
QY 110 LTAVAPADHTPPVDVDSRAILRRQNLSTPLTSSVATGTVLNYAAPLPLPLQDG 169
DB 3233 MSTATPS-STPE-----TWHTSTVLTTTATTTRGVSATPSS-----TPG 3271
QY 170 T--NTHIMATEASNAQYRVARATIRYRPLVPNAVGGYALISGFWPOTTITPT----- 220
DB 3272 TAHTTKVPITTTTGFT-----ATPSSSPGALTTPPWISTTTTPTTRGSTVT 3318
QY 221 --SVDNNSITSDVRLVQPGIASELVPSERLHYRNQGRWSVETSGVAEEATSGLVML 278
DB 3319 PSSIPGTHTATVLTITTTTAVATGSMATPSS-----STQTSPTPSLTTATTIT 3368
QY 279 CIHGLSVN-----SVTNPVTGALG-----LLDPALELEF 308
DB 3369 AT--GSTNPSSTGCTPIPPVLTTTATPAATSGVTTP--SALGTHTPPVPNTATHG 3426
QY 309 RNLTPGNTNTRVRYSTSTARHLRGADGTAEIT-----TTAAT----- 347
DB 3427 RSLPPSPHTVPTAWTSAT-----SGILGTHITEPSTGSHTPAATGTTQSTPALSS 3481
QY 348 -----RPMKDLTYTSNGVGEIORGIALTLFNL-----ADTLGLGLPTE-----LISSAG 392
DB 3482 PHSRSTTESPPSPGTTTTPGHT--RGTSRKTATATPSTKRTSTLLPSSPTSAPITVVVTG 3540
QY 393 GO-----LFYSRPVWSANGEPVKLYTSVE-----NAQDQKGI----- 425
DB 3541 CEFQCAWSEWLDYSPWPGPSG--GDFDTYSNIRAAGGAVCEQPLGLECRQAQGVPLRE 3599
QY 426 ---AIPHDIDLG-----ESRVVIQDYNQHQDRP-----TPSPAP 458
```



```
Db 562 TFDSTSVNLNSGLIGIYGQTSIECTSPSSVNVSTTKDGACFTKSVMPRLGGTYP 621
Qy 386 ELISSAGGOLFYSRPPVSVSANGPTVKLYTSVENAQDKGIAIPHDIDLGESVWIQYDN 445
Db 622 STFGVPGNTF--RATMTTDDK--KVYTYAN-----VYIQEYSS 657
Qy 446 QH-EQDRPT-----PSPAPSPFPFVLRANDVLWLSLTAAYDQSTYSGSTGPVYVS--- 495
Db 658 TTIESSESTSAVASSTSPSPSTLSTSTVTFPSSTRESSDSTTSAGSTTLQESITTT 717
Qy 496 --DSVTLNVNATGAQAVASLWTKVTLDRPLSTIQYSKTFVL 539
Db 718 SEESTTDSSTTATSTSTSSPSTADSTSTLSVDQFD---FIL 760

RESULT 12
YAG3_YEAST
ID YAG3_YEAST STANDARD; PRT; 1322 AA.
AC P39712;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Hypothetical 138.1 kDa protein in FLO9-GDH3 intergenic precursor.
GN YAL063C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=95249563; PubMed=7731988;
RA Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,
RA Hall J., Ouellette B.F.F., Keng T., Barton A.B., Su Y., Davies C.K.,
RA Storms R.K.;
RT "The nucleotide sequence of chromosome I from Saccharomyces
RT cerevisiae."
RL Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995).
CC -!- SIMILARITY: STRONG, TO YEAST PROTEIN FLO1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U12980; AAC04971.1; -.
DR GernOnline; 138402; -.
DR SGD; S0000059; YAL063C.
DR InterPro; IPR001389; Flocculin.
DR Pfam; PF00624; Flocculin; 13.
KW Hypothetical protein; Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 24
FT CHAIN 25 1322 HYPOTHETICAL PROTEIN YAL063C.
FT TAG 366 388
FT TRANSMEM 754 775
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 509 509 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 554 554 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 599 599 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 644 644 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 689 689 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 734 734 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 888 888 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1322 AA; 138072 MW; AADFDF1F13267CEA CRC64;

Query Match 3.8%; Score 131; DB 1; Length 1322;
Best Local Similarity 19.2%; Pred. No. 1.7; Mismatches 289; Indels 166; Gaps 26;
Matches 130; Conservative 92;

Qy 42 WCDRVDSQPFALPYIHPTNPFAPDVTAAAGAGPRVQRPALGSAWRDQAQRP----- 95
Db 643 WNDTFTSISTEMTTVTGNGLTDETIIVIRPTTATTMTTQPNWDTFTSTSEITTV 702
Qy 96 -----VASRRRPTTAGAAPLTAAPADHTPPVPDVSRSAGAILRRQYNLSPLT 144
Db 703 TCTNGLPTDETIIVIRPTTATTA-MTTTQPNWT-----FTSTSTEMTIVTCT 750
Qy 145 SSVATGTLNLYAAPLS-----PLLPLODGTNTHIMATEASNYAQYVARATIRYR 195
Db 751 NGVPTDETIVIRPTTSEGLISTTTEPTMTGTFTSTEMTTVTGNGQPTDETVIVIR-- 808
Qy 196 PLVPNAVGGVAISISFWPOTTTTPTSDMNSITSD-----VRILVOP--GIASELVI 246
Db 809 --TPTSEGLVTTTTEPTWGTFTS--TSTEMTITGNGQPTDETVIVIRPTTATSSLS 865
Qy 247 PSERL-----HYRQGRSVETSGVAEEATSGLMVLCIHGSLV-----N 286
Db 866 SSGQITSFITSARPIITPPPSNGTSVISSSVSSDTSLSVSSVTSSLVSSPVISS 925
Qy 287 SYNTPY-----TGALGLDFALEFRNLTPGNTNRVRSYSTARHLRRGADGTAEIT 342
Db 926 SPISSPVISSITTSASILS---ESSKSVIPTSSSTSGSSEST-----GSASSASS 975
Qy 343 TTAATREFMKDLY-----FTSTNGVGEIRG--IALTFLNADTL-----GGLPTE 386
Db 976 SSISSSPKSYSSSLPPVTSATTSQBITSSLPPVTITKTSEQTTLVTVTSCSHVCTE 1035
Qy 387 LISSAGGOLFYSRPPVSVSANGPTVKLYT-----SVENAQDKGIAIPHDIDL 433
Db 1036 SISSA-----IVSTATVTVSGATTE--YTTWCPSTITEITKQTETTKQKG--TTEQTET 1088
Qy 434 GSRVVIQYDYNQHQDRTPSPA-----PSRPFVLRANDVLWLSLTAAYEQST----- 484
Db 1089 TQTTTWTVISSCESDVCSTASPAIVSTSTATNGVTTEYTTWCPSTISTTESKQOTTLTV 1148
Qy 485 -----YGSSNGPVVSDSVLVNVAUQAQAVASLWTKVTLDRPLSTIQYSKTF 537
Db 1149 TSCGGVCSSETSPAIVSTATATVN-----DVVTVYSTWRPDTTNEQ----- 1190
Qy 538 VLPLRGKLSFWBAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLG-AGFVSIS 596
Db 1191 --SVSSKMSATSEITT-----NTGAAETTTTGAETKTVTSSISRFNHAETQTAS 1241
Qy 597 AVAVLAPHSALELLEDT 613
Db 1242 ATDVIGHSSSVSVSET 1258

RESULT 13
ID AMYI_SACDI STANDARD; PRT; 768 AA.
AC P29760;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glucoamylase S2 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)
DE (1,4-alpha-D-glucan glucohydrolase) (GAII).
GN STA2 OR DEX1.
OS Saccharomyces diastaticus (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=41870;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94288586; PubMed=8017901;
```


RA Kim K., Bajszar G., Lee S.Y., Knudsen F., Mattoon J.R.;
 RT "Cloning of a new allelic variant of a Saccharomyces diastaticus
 RT glucosylase gene and its introduction into industrial yeasts."; Appl. Biochem. Biotechnol. 44:161-195(1994).
 RL [2]
 RN
 RP
 RX
 RX MEDLINE=91276266; PubMed=2055484;
 RA Lambrechts M.G., Pretorius I.S., Sollitt P., Marmur J.;
 RT "Primary structure and regulation of a glucosylase-encoding gene
 RT (STA2) in Saccharomyces diastaticus."; Gene 100:95-103(1991).
 RL
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M60650; AAA35107.1; -;
 DR EMBL; M90490; AAA20560.1; -;
 DR PIR; J0474; J0474.
 DR HSP; P08017; IAYX.
 DR InterPro; IPR00165; Glyco hydro 15.
 DR InterPro; IPR00828; Glyco trans 6hp.
 DR Pfam; PF00723; Glyco hydro 15; 1.
 DR PRINTS; PR00736; GLHYDRLASE15.
 DR PROSITE; PS00820; GLUCOAMYLASE; 1.
 DR Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KW Signal; Multigene family.
 FT SIGNAL 1 21
 FT CHAIN 22 768 GLUCOAMYLASE S2.
 FT DOMAIN 22 348 SER/THR-RICH.
 FT DOMAIN 349 692 H SUBUNIT.
 FT DOMAIN 693 768 Y SUBUNIT.
 FT BINDING 456 456
 FT ACT SITE 519 519
 FT ACT SITE 522 522 CATALYTIC BASE (BY SIMILARITY).
 FT ACT SITE 523 523 GENERAL ACID CATALYST (BY SIMILARITY).
 FT CARBOHYD 35 35 INTERACT WITH SUBSTRATES (BY SIMILARITY).
 FT CARBOHYD 309 309 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 424 424 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 435 435 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 514 514 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 651 651 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 721 721 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 742 742 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 164 164 MISSING (IN REF. 1).
 FT CONFLICT 624 624 D -> N (IN REF. 1).
 SQ SEQUENCE 768 AA; 82586 MW; 3FAC172C129A0C6F CRC64;
 Query Match 3.8%; Score 130.5; DB 1; Length 768;
 Best Local Similarity 20.8%; Pred. No. 0.96;
 Matches 130; Conservative 70; Mismatches 231; Indels 193; Gaps 29;
 QY 4 RPILL-----LLMLPMLPAPPQPSRRRRRRSGSGGFGWDRVDSQPPFAIPYHPT 59
 DB 3 RPFLLAYLVLSLNFNSALGFPTALVP-----RGSSS-----SNITSSGPPSTPFSAT 50
 QY 60 NPAPDVTAAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAALPLTAVAPADHT 119
 DB 51 ESRSTGTVT-----FSSSKYPSKSTSTSVSTTTTIVTTTSTTTTSTTTSTTT 98
 QY 120 PPVPDVSRRGAILRRQYNLSTSP--LTSSVATGTN---LVLYAAPLSPPLPLQDGTNTHI 174

DB 99 TITTTVCSGTNSAGETSCGSKPTITTTVPCSTSTSESTTTTPTTPTTVVSTTV 158
 QY 175 MATEASNYAQYRVARATIRYRPLVNAVAGYAI-----SWPQT-TTTPTSVD 223
 DB 159 VTTEYASTSTKQGGIITTF--VTKNPTTYLTATPTSSVTVTNFTTTTITTVCGSTG 216
 QY 224 MNSITSDVRIIVQPGIASLVIPSRHLHYRNQGRSVETSGVAEEBETSLVLMCLHGS 283
 DB 217 TNSAGET-----TSGCSPKTVTT--TVPCTGT 242
 QY 284 LVNSYT---NTPYTGALGLLDFALEFRNLTPQNTNTRVSRYSSTARHLRRGADGTAE 340
 DB 243 -GEYITEAPVTAV-----TTTVTTESST-----GTNSAGK 275
 QY 341 LTTAATRRMKDLYFTSTNGVGEIGRGIALTLNLAADTLGLLPTELISAGGQLFYSRP 400
 DB 276 TTTSYTTKSPVTTY-----VDFGKGI-----LDQSCGG---VFSNNGSSQVQLRD 318
 QY 401 VVSANGEPTVKLYTS-----VENAQODKGIAIPHDID-LGESRVVTQDYDNQHEQD 450
 DB 319 VVLWNG--TV-VYDSNGAMDSSPLEEWLQKQKVSIERIFENIGPSAVY----- 364
 QY 451 RPT-----PSPAPSRP-----FVLRANDVLMLSLTA-----EYD 481
 DB 365 -PSILPGWVIASPSQTHPDYFYQWIRDSALTINSIVSHSADPAETLLQYLVNVSFHLQRT 423
 QY 482 QSTYGSSTGPVYVSDSVTLNVATGAQAVARSIDWTKVLDGRPLSTIQOYKTFVPL 541
 DB 424 NTLGAGIG--YNTDVTALGDPKNNVDNTAFTBPWGRPQNDGPALRSIA-----ILKI 474
 QY 542 RGKLSFEWAGTTKAGYPYNYNTTA 565
 DB 475 ---IDYIKSGTDLGAKYPPQSTA 495
 RESULT 14
 DAN4 YEAST
 ID DAN4 YEAST STANDARD; PRT; 1161 AA.
 AC P47179;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cell wall protein DAN4 precursor.
 GN DAN4 OR YJR151C OR J2223.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Scareez T.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 RP REGULATION.
 RX MEDLINE=21113168; PubMed=11160904;
 RA Cohen B.D., Sertil O., Abramova N.E., Davies K.J., Lowry C.V.;
 RT "Induction and repression of DAN1 and the family of anaerobic
 RT mannoprotein genes in Saccharomyces cerevisiae occurs through a
 RT complex array of regulatory sites."; Nucleic Acids Res. 29:799-808(2001).
 RL Nucleic Acids Res. 29:799-808(2001).
 CC -1- FUNCTION: Component of the cell wall (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Potential).
 CC -1- PTM: Extensively O-glycosylated (Potential).
 CC -1- SIMILARITY: Belongs to the SRP1 / Tip1 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```
CC or send an email to license@isb-sib.ch).
CC EMBL; Z49651; CA489684.1;
CC PIR; S57180; S57180.
CC GERMOnline; 141984;
CC SGO; S0003912; DAN4.
CC GO; GO:0009277; C:cell wall (sensu Fungi); IDA.
CC DR Pfam; PF00660; SRP1_TIP1.
CC DR PROSITE; PS00724; SRP1_TIP1; 1.
CC KW Cell wall; Glycoprotein; Membrane; GPI-anchor; Signal; Lipoprotein.
CC FT SIGNAL 1 24
CC FT CHAIN 25 1146
CC FT PROPEP 1147 1161
CC FT LIPID 1146 1146
CC FT SEQUENCE 1161 AA; 118358 MW; 7954C15D69F0CA58 CRC64;
CC Query Match 3.8%; Score 130.5; DB 1; Length 1161;
CC Best Local Similarity 17.2%; Pred. No. 1.5;
CC Matches 104; Conservative 95; Mismatches 274; Indels 133; Gaps 21;
CC QY 57 HPTNPAPDVTAAG-----AGPRVQRPARPLGSHWRDQARPAVASRRR--- 101
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 59 HKITETPSEIAAAVDYGDFTRLTGISGDEVTMRI--TGVFWYSTRLKPATSSALSKDG 116
CC QY 102 -----PTTAGAAPLTAAPAHDTTPVDVDSRGAILRQYNNLSTPLTSSVA---TGTN 152
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 117 IYALPTSTSTTTKSSSTSTPTTITSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 176
CC QY 153 LVLVAAPLPLDQGTNTHTIMATEAGNYAQRVARATIRYRLVFNPAVAGYALISFW 212
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 177 STPTTSTSTPTTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 224
CC QY 213 POTTTPPSVDMNSTSTDRILVOPGIASELVPISERLHYRNQGRSVETSGVAEEBAT 272
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 225 PTTSTPTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 283
CC QY 273 SGLVMLCIHGLSVNVTNTPYTGALGLD--FALELEFRLNTPGNTNTRVSRYSSTARH 330
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 284 TSTTSTAPTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 340
CC QY 331 LRRGADGTAEITTAATRFMKDLYFTSNGVGEIGRGIALTLFNLAIDLGLLPTELSS 390
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 341 ----VSTINAFTSATTTTSTSTYSSSS-----PSQVTSS 372
CC QY 391 AGGQLFYRPRVVSANGBPTV--KLTVSVEAQQDQKGIAPHIDIDGEKRVVIQDYNDQHE 448
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 373 A-----EPTTVSEVTSSVEPTRSSQ-----VTSSAEPTTVSEPTSSVE 410
CC QY 449 QDRPTFSPAPRPFVSLRNDVLMLSLTAAYDQSTYGSSTGTGVVSDSVTLVNVATGQ 508
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 411 PTRSQVTSSEAP-----TTVSEFTSSVEPTRSSQVTSSEAPTTVSEFTSSV 457
CC QY 509 AVARSLDKTKVTLGDRPLSTIQYQSKTFVFLPLR--GKLSFWEAGTTKAGYPVNYNTASD 567
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 458 EPTRS---SQVTSSEAP--TTVSEFTSS--VEPTRSSQVTSSEAPTTVSEFTSSVEPTRSS 511
CC QY 568 QLLVNEAGHRAVIASTYTTSLGAGPVISAVAVLAPHALALLEDTLDPARAHTFDFFC 627
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 512 QV--TSSAEPTTVSEVTSSV--EPIRSSQVTTTEPVSSF-----GSTFSEIT 554
CC QY 628 PECREPL 633
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 555 SSAEPL 560
CC RESULT 15
CC PCGN HUMAN
CC ID PCGN HUMAN STANDARD; PRT; 1321 AA.
CC AC O14594; Q9UPK6;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
```

```
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3).
GN CSPG3 OR NCAN OR NEUR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99013874; PubMed=9795216;
RA Prange C.K., Pennacchio L.A., Lieuallen K., Fan W., Lennon G.G.;
RT "Characterization of the human neurocan gene, CSPG3.";
RL Gene 221:199-205(1998).
RN [2]
RP SEQUENCE OF 1-990 AND 1007-1321 FROM N.A., AND VARIANT ALA-1254.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stiilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,
RA Dangnan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Krommiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RT "Sequence analysis of an ~1 Mb region containing the MEF2B gene in
RT 19p12.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
CC development by binding to neural cell adhesion molecules (NC-CAM
CC and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
CC acid.
CC -!- TISSUE SPECIFICITY: Brain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF026547; AAC80576.1; -.
CC EMBL; AC003110; AAB86655.1; -.
CC EMBL; AC005254; AAC25581.1; -.
CC HSP; P00740; 1EDM.
CC Genew; HGNC:2465; CSPG3.
CC MIM; 600826; -.
CC InterPro; IPR000152; Asx hydroxyl_s.
CC InterPro; IPR000742; EGF 2.
CC InterPro; IPR001881; EGF Ca.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; Xlink; 2.
CC PRINTS; PR01265; LINKMODULE.
CC ProDom; PD000918; Link; 2.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00409; IG; 1.
CC SMART; SM00445; LINK; 2.
```

DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS00041; C-TYPE LECTIN 2; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal; Polymorphism.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 1321 NEUROCAN CORE PROTEIN.
FT DOMAIN 38 153 IG-LIKE V-TYPE.
FT DOMAIN 159 254 LINK 1.
FT DOMAIN 260 356 LINK 2.
FT DOMAIN 1008 1044 EGF-LIKE 1.
FT DOMAIN 1046 1082 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1084 1213 C-TYPE LECTIN.
FT DOMAIN 1214 1272 SUSHI.
FT DISULFID 59 140 BY SIMILARITY.
FT DISULFID 182 253 BY SIMILARITY.
FT DISULFID 206 227 BY SIMILARITY.
FT DISULFID 280 355 BY SIMILARITY.
FT DISULFID 304 325 BY SIMILARITY.
FT DISULFID 1012 1023 BY SIMILARITY.
FT DISULFID 1017 1032 BY SIMILARITY.
FT DISULFID 1034 1043 BY SIMILARITY.
FT DISULFID 1088 1099 BY SIMILARITY.
FT DISULFID 1116 1208 BY SIMILARITY.
FT DISULFID 1184 1200 BY SIMILARITY.
FT DISULFID 1215 1258 BY SIMILARITY.
FT DISULFID 1244 1271 BY SIMILARITY.
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1026 1026 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1223 1223 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 1254 1254 V -> A (in dbSNP:1064389).
FT FT /FTID=VAR_016176.
FT CONFLICT 1234 1234 N -> Y (IN REF. 2).
FT CONFLICT 1282 1282 G -> R (IN REF. 2).
SQ SEQUENCE 1321 AA; 142972 MW; 2EF47F823DB980B8 CRC64;

Query Match 3.8%; Score 129.5; DB 1; Length 1321;
Best Local Similarity 22.0%; Pred. No. 2; Indels 201; Gaps 32;
Matches 140; Conservative 63; Mismatches 231

QY 22 PGQSGRRRRGRRSGSGGF-----WGDRVDSQP-----FAIPYIHPTNPFAPDVT 67
DB 472 PTDPMERRRRGFKLNGRYFQQQPEPEGLQGGMEASQPPTSBAVYQMEP--PLAMAVT 529
QY 68 AAGAGPRVROPAPRLGSAWRD---QAORPAVASRRRPTTAGAA-----PLTAVAPA 116
DB 530 EMLGSG-QSRSP-----WADLTNEYDMFGAGS-----AGGKSPEPWLWPTMVPVS 575
QY 117 ---HDTPPVPDVP-SRGAILR-RQYNLTSPLTSSVATGTLNLYAAPLS-----PLLPL 166
DB 576 ISGHSRAPULELEKAGPSARPATPDLFWSPLEATVSAPAPWEAFPPVATSDPLPMAM 635
QY 167 QDGTNTTHIM-----ATEASNAQYRVARATIRYRPLVPNAVGGYAIISIFWPTTTTPT 220
DB 636 LRGPKEWMLPHPTPISTEANRVEAHGEATATAPSPAETKV--YSLPLSL-----TPT 687
QY 221 SVMDSITSDVILLVQPGIASLIVPSERLHYRNQWRVSFVTSVGAEEAETS--GLVML 278
DB 688 G-----QGGEAMPTTPESPRADFRETGETSPAQVNKAHSSSSPWPVSNR 732
QY 279 CIHGLSVNTPTPYTGALGLDFALEPRNLTPGNTNTRVSRYSSTARHRLRRGADGT 338
DB 733 NVAVGFPVPTETATEPTGLRGI-----PGSES-----GVFDT 763
QY 339 AEUTTTAATRFMKDLY-----FTSTNGVGEIGR-GIALT----- 371

Db 764 AESPTSGLQATVDEVDQWPFSVYVSKGLDASSPSAPLPGVFLVPKVTNPNLPWATDEG 823
QY 372 -LFNLADTLLGGLPTL---ISSAGGQLFYSRPVSANGPTVKLYTSV-----ENAOQDKG 424
Db 824 PTVNPMDSTVTFAPSDASGIWEPGQVF--EEAESTTLPQVALDTSIVTPLTTLEQGDK 881
QY 425 IAIPHDIDILGESRVVIQDYDNQHEQDRPTPSP-----APSRPFSVLRANDVL 471
Db 882 VGVPAMSTLGSS-----SSQHPPEPDQVETQGTSGASVPPHQSSPLGKEPVP 929
QY 472 WLSLTAAYDQSTYSGSTGPVYVD-SVTLVNVTGAQAVARSLDWTKVTLDCRPLSTIQ 530
Db 930 PGFTPTAASVGESASVSSGPTVPWDPSSILLPVTLGIE-----DPELEVLAGSP----- 978
QY 531 QYSKTFVFLPLRKLKLSFWE---AGITKA--GYPN 560
Db 979 -----GVESFWEVASGEEPALPGTPMN 1001

Search completed: August 16, 2004, 13:34:13
Job time : 17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 13:24:37 ; Search time 42 seconds

(without alignments)
4958.143 Million cell updates/sec

Title: US-09-851-410A-8

Perfect score: 3408

Sequence: 1 MRPRILLLLMLPMLPAP.....QSTVAELQLKMKVGTREL 660

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	3392	99.5	660	12 Q9E8G5	Q9E8G5 hepatitis e
2	3385	99.3	660	12 Q91855	Q91855 hepatitis e
3	3383	99.3	660	12 Q9WQAO	Q9WQAO hepatitis e
4	3382	99.2	660	12 Q89468	Q89468 hepatitis e
5	3382	99.2	660	12 Q81871	Q81871 hepatitis e
6	3381	99.2	660	12 Q9W7W9	Q9W7W9 hepatitis e
7	3374	99.0	660	12 Q89419	Q89419 hepatitis e
8	3372	98.9	660	12 Q81878	Q81878 hepatitis e
9	3367	98.8	660	12 Q91856	Q91856 hepatitis e
10	3367	98.8	660	12 Q89411	Q89411 hepatitis e
11	3362	98.7	660	12 Q9WL14	Q9WL14 hepatitis e
12	3342.5	98.1	659	12 Q8JT32	Q8JT32 hepatitis e
13	3336	97.9	660	12 Q88985	Q88985 hepatitis e
14	3379	93.3	674	12 Q8JUN6	Q8JUN6 hepatitis e
15	3176	93.2	660	12 Q8JUN8	Q8JUN8 hepatitis e
16	3175	93.2	660	12 Q91114	Q91114 hepatitis e

17	3174	93.1	660	12 Q8JUN4	Q8JUN4 hepatitis e
18	3174	93.1	660	12 Q8AZM3	Q8AZM3 hepatitis e
19	3172	93.1	660	12 Q9YLQ9	Q9YLQ9 hepatitis e
20	3171	93.0	660	12 Q9YLK2	Q9YLK2 hepatitis e
21	3167	92.9	671	12 Q8JUN1	Q8JUN1 hepatitis e
22	3166	92.9	660	12 Q8V729	Q8V729 swine hepat
23	3166	92.9	674	12 Q8JUN2	Q8JUN2 hepatitis e
24	3163	92.8	660	12 Q36613	Q36613 swine hepat
25	3160	92.7	674	12 Q806E0	Q806E0 swine hepat
26	3160	92.7	674	12 Q806D7	Q806D7 hepatitis e
27	3160	92.7	674	12 Q7TGF1	Q7TGF1 hepatitis e
28	3159	92.7	660	12 Q8BB05	Q8BB05 swine hepat
29	3158	92.7	674	12 Q80IR5	Q80IR5 hepatitis e
30	3157	92.6	660	12 Q80IY5	Q80IY5 hepatitis e
31	3153	92.5	660	12 Q80IR7	Q80IR7 hepatitis e
32	3103	91.1	672	12 Q9IVZ8	Q9IVZ8 hepatitis e
33	3053	89.6	605	12 Q9WLO0	Q9WLO0 hepatitis e
34	2688	78.9	550	12 Q8JVV3	Q8JVV3 hepatitis e
35	2660	78.1	525	12 Q39947	Q39947 hepatitis e
36	2381	69.9	486	12 Q9WNN2	Q9WNN2 hepatitis e
37	2298	67.4	466	12 Q9WNN1	Q9WNN1 hepatitis e
38	2197	64.5	436	12 Q9WI48	Q9WI48 hepatitis e
39	1762	51.7	344	12 Q9WJZ5	Q9WJZ5 hepatitis e
40	1416	41.5	286	12 Q8JQ92	Q8JQ92 swine hepat
41	1411	41.4	286	12 Q8JPR3	Q8JPR3 swine hepat
42	1409	41.3	286	12 Q8JQ93	Q8JQ93 swine hepat
43	1405	41.2	283	12 Q9YEB1	Q9YEB1 hepatitis e
44	1334.5	39.2	606	12 Q913Y7	Q913Y7 avian hepat
45	1320	38.7	258	12 Q81861	Q81861 hepatitis e

ALIGNMENTS

RESULT 1

Q9E8G5 PRELIMINARY; PRT; 660 AA.

AC Q9E8G5; DT 01-NAR-2001 (TREMBLrel. 16, Created)

DT 01-NAR-2001 (TREMBLrel. 16, Last sequence update)

DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Structural protein.

OS Hepatitis E virus.

OC Viruses: ssRNA positive-strand viruses, no DNA stage;

OC Hepatitis E-like viruses.

OX NCBI_TaxID=12461;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Abb-2B;

RX MEDLINE=20271579; PubMed=10813471;

RA van Cuyck-Gandre H., Zhang H.Y., Tsarev S.A., Warren R.L.,

RA Caudill J.D., Snellings N.J., Begot L., Innis B.L., Longer C.F.;

RT "Phylogenetically distinct hepatitis E viruses in Pakistan.";

RL Am. J. Trop. Med. Hyg. 62:187-189(2000).

DR EMBL; AF185822; AAG16766.1; "

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR004261; SP2.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF03014; SP2; 1.

SQ SEQUENCE 660 AA; 70903 MW; 1F506BE3CFB3BACE CRC64;

Query Match 99.5%; Score 3392; DB 12; Length 660;

Best Local Similarity 99.5%; Pred. No. 1.2e-213;

Matches 657; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MRPRILLLLMLPMLPAPPPGSGRRRRSGSGGFGWGDVDSQPPAIPYIHTN 60

Db 1 MRPRILLLLMLPMLPAPPPGSGRRRRSGSGGFGWGDVDSQPPAIPYIHTN 60

Qy 61 PFAPDVTAAAGAGPRVRQPARPLGSAWMDQQRPAVASRRRPTTAGAAPLTAAPAHDT 120

Db 61 PFAPDVTAAAGAGPRVRQPARPLGSAWMDQQRPAVASRRRPTTAGAAPLTAAPAHDT 120

```
QY 121 PVPDVSRAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAISISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 240
Db 181 NYAQYVARATIRYRPLVNAVGGYAISISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 240
QY 241 ASLVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSVNSYNTPTYTGAIGLL 300
Db 241 ASLVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSVNSYNTPTYTGAIGLL 300
QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGBIGRGIALTLFNADTLTLLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGBIGRGIALTLFNADTLTLLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSPFVLRANDVLMVLSLTAAY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSPFVLRANDVLMVLSLTAAY 480
QY 481 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQOYKTFVFLP 540
Db 481 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQOYKTFVFLP 540
QY 541 LRKGLSFWAGTTKAGYPYNYNTTASDQLLVENAAAGHVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWAGTTKAGYPYNYNTTASDQLLVENAAAGHVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDLTDYPARAHTFDDFCPCRLGLOGCAFOSTVAELQRLKMKVGKTREL 660
Db 601 LAPHSALELLEDLTDYPARAHTFDDFCPCRLGLOGCAFOSTVAELQRLKMKVGKTREL 660
```

RESULT 2

```
091855 PRELIMINARY; PRT; 660 AA.
ID 091855
AC 091855;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=93-Egypt;
RX MEDLINE=99105430; PubMed=9890424;
RA Tsarev S.A., Binn L.N., Gomatos P.J., Arthur R.R., Monier M.K.,
van Cuyck-Gandre H., Longer C.F., Innis B.L.;
RT "Phylogenetic analysis of hepatitis E virus isolates from Egypt.";
RL J. Med. Virol. 57:68-74(1999).
DR ENBL; AF051351; AAC35761.1; "-.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71040 MW; 044FF5EA7C492791 CRC64;
```

Query Match 99.3%; Score 3385; DB 12; Length 660;
Best Local Similarity 99.1%; Pred. No. 3,4e-213;
Matches 654; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 MRPRPILLLMLFPLPMLPAPPQPSGRRRSGSGGFGWGDVDSQFPAIPIHTPTN 60
Db 1 MRPRPILLLMLFPLPMLPAPPQPSGRRRSGSGGFGWGDVDSQFPAIPIHTPTN 60
QY 61 PFAPDVTAAAGPRVRQPARPLGSAWRDQAQRPAVARRRPTTAGAAPTAVAPADHTP 120
```

```
Db 61 PFAPDVTAAAGPRVRQPARPLGSAWRDQAQRPAVARRRPTTAGAAPTAVAPADHTP 120
QY 121 PVPDVSRAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAISISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 240
Db 181 NYAQYVARATIRYRPLVNAVGGYAISISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 240
QY 241 ASLVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSVNSYNTPTYTGAIGLL 300
Db 241 ASLVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSVNSYNTPTYTGAIGLL 300
QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGBIGRGIALTLFNADTLTLLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGBIGRGIALTLFNADTLTLLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSPFVLRANDVLMVLSLTAAY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSPFVLRANDVLMVLSLTAAY 480
QY 481 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQOYKTFVFLP 540
Db 481 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQOYKTFVFLP 540
QY 541 LRKGLSFWAGTTKAGYPYNYNTTASDQLLVENAAAGHVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWAGTTKAGYPYNYNTTASDQLLVENAAAGHVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDLTDYPARAHTFDDFCPCRLGLOGCAFOSTVAELQRLKMKVGKTREL 660
Db 601 LAPHSALELLEDLTDYPARAHTFDDFCPCRLGLOGCAFOSTVAELQRLKMKVGKTREL 660
```

RESULT 3

```
09WQAO PRELIMINARY; PRT; 660 AA.
ID 09WQAO
AC 09WQAO;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF-2.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AKL-90;
RX MEDLINE=99350000; PubMed=10423137;
RA Arankalle V.A., Paranjape S., Emerson S.U., Purcell R.H..
RA Walimbe A.M.;
RT "Phylogenetic analysis of hepatitis E virus isolates from India (1976-1993).";
RL J. Gen. Virol. 80:1691-1700(1999).
RL J. Gen. Virol. 80:1691-1700(1999).
DR EMBL; AF124407; AAD45493.1; "-.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71041 MW; 03B72DDF0AB7B521 CRC64;
```

Query Match 99.3%; Score 3383; DB 12; Length 660;
Best Local Similarity 99.2%; Pred. No. 4,6e-213;
Matches 655; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 MRPRPILLLMLFPLPMLPAPPQPSGRRRSGSGGFGWGDVDSQFPAIPIHTPTN 60
```

```
Db 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60
Qy 61 PFAPDVTAAAGPVRQPARPLGSAWRDQAPAVASRRRTTGAAPLTAAPAHDTTP 120
Db 61 PFAPDVTAAAGPVRQPARPLGSAWRDQAPAVASRRRTTGAAPLTAAPAHDTTP 120
Qy 121 PVPDVSAGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSAGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Qy 181 NYAQYRVARATIRYRPLVPNAVGGYVAISISFWPQTTTTPTSDVMNSITSTDVRIILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVPNAVGGYVAISISFWPQTTTTPTSDVMNSITSTDVRIILVQPGI 240
Qy 241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTNTPTTGALGLL 300
Db 241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTNTPTTGALGLL 300
Qy 301 DFALIEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTTAAATREMKDLYFTSTNG 360
Db 301 DFALIEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTTAAATREMKDLYFTSTNG 360
Qy 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Qy 421 QDKGIAIPHDIDLGESRVVIQDYNQHQEDRTPSPAPSRPFSVLRANDVLSLTAAY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYNQHQEDRTPSPAPSRPFSVLRANDVLSLTAAY 480
Qy 481 DOSTYGSSTGPPYVSDSVTLNVNATCAQAVASRLDWTKTLDGRPLSTTQQYKTFVFLP 540
Db 481 DOSTYGSSTGPPYVSDSVTLNVNATCAQAVASRLDWTKTLDGRPLSTTQQYKTFVFLP 540
Qy 541 LRKLSFWEAGTTKAGYPYNNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKLSFWEAGTTKAGYPYNNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Qy 601 LAPHASALLEDTLDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKREL 660
Db 601 LAPHASALLEDTLDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKREL 660

RESULT 4
Q89468 PRELIMINARY; PRT; 660 AA.
AC Q89468;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE UNNAMED protein product.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]_
RP SEQUENCE FROM N.A.
RA Yin S.R., Purcell R.H., Emerson S.U.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=X52-87;
RX MEDLINE=95176571; PubMed=7871758;
RA Yin S., Purcell R.H., Emerson S.U.;
RT "A new Chinese isolate of hepatitis E virus: comparison with strains
RT recovered from different geographical regions.";
RL Virus Genes 9:23-32(1994).
DR EMBL; L25547; AAA31080.1; -.
DR EMBL; L25595; AAA65490.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
```

```
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70979 MW; DC68116DCD639175 CRC64;

Query Match 99.2%; Score 3382; DB 12; Length 660;
Best Local Similarity 99.1%; Pred. No. 5.3e-213;
Matches 654; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60
Db 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60
Qy 61 PFAPDVTAAAGPVRQPARPLGSAWRDQAPAVASRRRTTGAAPLTAAPAHDTTP 120
Db 61 PFAPDVTAAAGPVRQPARPLGSAWRDQAPAVASRRRTTGAAPLTAAPAHDTTP 120
Qy 121 PVPDVSAGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSAGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Qy 181 NYAQYRVARATIRYRPLVPNAVGGYVAISISFWPQTTTTPTSDVMNSITSTDVRIILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVPNAVGGYVAISISFWPQTTTTPTSDVMNSITSTDVRIILVQPGI 240
Qy 241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTNTPTTGALGLL 300
Db 241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTNTPTTGALGLL 300
Qy 301 DFALIEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTTAAATREMKDLYFTSTNG 360
Db 301 DFALIEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTTAAATREMKDLYFTSTNG 360
Qy 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Qy 421 QDKGIAIPHDIDLGESRVVIQDYNQHQEDRTPSPAPSRPFSVLRANDVLSLTAAY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYNQHQEDRTPSPAPSRPFSVLRANDVLSLTAAY 480
Qy 481 DOSTYGSSTGPPYVSDSVTLNVNATCAQAVASRLDWTKTLDGRPLSTTQQYKTFVFLP 540
Db 481 DOSTYGSSTGPPYVSDSVTLNVNATCAQAVASRLDWTKTLDGRPLSTTQQYKTFVFLP 540
Qy 541 LRKLSFWEAGTTKAGYPYNNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKLSFWEAGTTKAGYPYNNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Qy 601 LAPHASALLEDTLDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKREL 660
Db 601 LAPHASALLEDTLDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKREL 660

RESULT 5
Q81871 PRELIMINARY; PRT; 660 AA.
ID Q81871
AC Q81871;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE ORF 2 precursor.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
RN [2]
```

RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K., Min K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92335008; PubMed=1630924;
RA Aye T.T., Uchida T., Ma X.Z., Iida F., Shikata T., Zhuang H.,
RA Win K.M.;
RT "Complete nucleotide sequence of a hepatitis E virus isolated from the
RT Xinjiang epidemic (1986-1988) of China.";
RL Nucleic Acids Res. 20:3512-3512(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92271462; PubMed=1589964;
RA Fry K.E., Tam A.W., Smith M.W., Kim J.P., Luk K.-C., Young L.M.,
RA Piatak M., Feldman R.A., Yun K.Y., Purdy M.A., McCaustland K.A.,
RA Bradley D.W., Reyes G.R.;
RT "Hepatitis E virus (HEV): Strain variation in the nonstructural gene
RT region encoding consensus motifs for an RNA-dependent RNA polymerase
RT and an ATP/GTP binding site.";
RL Virus Genes 6:173-185(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93348763; PubMed=9346669;
RA Bi S.-L., Purdy M.A., McCaustland K.A., Margolis H.S., Bradley D.W.;
RT "The sequence of hepatitis E virus isolated directly from a single
RT source during an outbreak in China.";
RL Virus Res. 28:233-247(1993).
DR EMBL: L08816; AAA03191.1; -;
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
KW Signal.
FT SIGNAL.
FT SIGNAL.
RP SEQUENCE 1 19 POTENTIAL.
Query Match 99.2%; Score 3382; DB 12; Length 660;
Best Local Similarity 99.1%; Pred. No. 5.3e-213;
Matches 654; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPFAPYIHPN 60
DB 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPFAPYIHPN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSARWDOAQPAAVARRRPTTAGAAPTAVAPADTP 120
DB 61 PFAPDVTAAAGAGPRVQPARPLGSARWDOAQPAAVARRRPTTAGAAPTAVAPADTP 120
QY 121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLLPLODGTNTHIMATEAS 180
DB 121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVNAVGYAISISFWQTTTTTPTSDVMSITSDVRLVQPGI 240
DB 181 NYAQYRVARATIRYRPLVNAVGYAISISFWQTTTTTPTSDVMSITSDVRLVQPGI 240
QY 241 ASEVLIPSERLHYNQWRSVETSGVAEEATSGLVMLCHGSLVNSYNTPTGTGALGLL 300
DB 241 ASEVLIPSERLHYNQWRSVETSGVAEEATSGLVMLCHGSLVNSYNTPTGTGALGLL 300
QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360

DB 301 DFALLEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLNADTLGGLPTELISAGGOLFYSRPPVSVANGSPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLNADTLGGLPTELISAGGOLFYSRPPVSVANGSPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDIGESRVVIQDYDNOHQDRPTSPAPSRPFVLRANDVLMLSLTAAY 480
DB 421 QDKGIAIPHDIDIGESRVVIQDYDNOHQDRPTSPAPSRPFVLRANDVLMLSLTAAY 480
QY 481 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSILDWTKVLDGRPLSTIQOYSKTFVLP 540
DB 481 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSILDWTKVLDGRPLSTIQOYSKTFVLP 540
QY 541 LRKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRAVIASTYTTSLGAGPVSISAV 600
DB 541 LRKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRAVIASTYTTSLGAGPVSISAV 600
QY 601 LAPHSALALLEDLTDYPARAHTFDDFCPECRPLGLOGCAFQSTVAELQRLKMKVKGTREL 660
DB 601 LAPHSALALLEDLTDYPARAHTFDDFCPECRPLGLOGCAFQSTVAELQRLKMKVKGTREL 660
RESULT 6
Q9W7W9 PRELLMINARY; PRT; 660 AA.
AC Q9W7W9;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-OCT-2003 (TEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Structural protein 2 (Structural viral protein).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HEV-Morocco;
RX MEDLINE=99107414; PubMed=9892396;
RA Meng J., Cong M., Dai X., Pillot J., Purdy M.A., Fields H.A.,
RA Khudyakov Y.E.;
RT "Primary structure of open reading frame 2 and 3 of the hepatitis E
RT virus isolated from Morocco.";
RL J. Med. Virol. 57:126-133(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Morocco;
RA Chen G., Meng J.;
RT "Identification of the 5' Capped and 3' Complete Terminal Sequence of
RT the Hepatitis E virus Isolated From Morocco.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF065061; AAD10627.1; -;
DR EMBL: AY230202; AAO72992.1; -;
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70943 MW; 9719D30CCD932950 CRC64;
Query Match 99.2%; Score 3381; DB 12; Length 660;
Best Local Similarity 99.2%; Pred. No. 6.2e-213;
Matches 655; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPFAPYIHPN 60
DB 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPFAPYIHPN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSARWDOAQPAAVARRRPTTAGAAPTAVAPADTP 120
DB 61 PFAPDVTAAAGAGPRVQPARPLGSARWDOAQPAAVARRRPTTAGAAPTAVAPADTP 120
QY 121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLLPLODGTNTHIMATEAS 180


```

Db 121 PVPDVDSRGAILLRQYNLSTPLTSSVATGCTNLVLYAAPLSPLLPLQDGTNTHIMATEAS 180
QY 181 NVAQYRVARATIRYRLPVNAVGGYAISISFWPQTITPTSDVMNSITSTDVRLVQPGI 240
Db 181 NVAQYRVARATIRYRLPVNAVGGYAISISFWPQTITPTSDVMNSITSTDVRLVQPGI 240
QY 241 ASSELVIPSRLHYRNQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGTALGLL 300
Db 241 ASSELVIPSRLHYRNQWRSVETSGVAEEATSGLVMLCIHGSPVNSYNTPTGTALGLL 300
QY 301 DFALELEFNLPNGTNTVRSYSSSTARHRLRGADGTAELTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFNLPNGTNTVRSYSSSTARHRLRGADGTAELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDLGESRVVIQDYDQHQEDRPTSPAPSRPFSVLRANDVLMSLTAAEY 480
Db 421 QDKGIAIPHDIIDLGESRVVIQDYDQHQEDRPTSPAPSRPFSVLRANDVLMSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLQGRPLSTIQOYKSTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLQGRPLSTIQOYKSTFFVLP 540
QY 541 LRKGLSFWAGTTKAGYPYNYNTASDQLLVENAGHRVAISYTTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWAGTTKAGYPYNYNTASDQLLVENAGHRVAISYTTTSLGAGPVSISAVAV 600
QY 601 LAPHSALEDTLDYPARAHTFDDFCPCPRPLGQCAFQSTVABLQRLKMKVGKTRREL 660
Db 601 LAPHSALEDTLDYPARAHTFDDFCPCPRPLGQCAFQSTVABLQRLKMKVGKTRREL 660

RESULT 7
Q69419
ID Q69419 PRELIMINARY; PRT; 660 AA.
AC Q69419;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE ORF1, ORF2 & ORF3.
OS Hepatitis E virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RA von Brunn A., Seebach J., Thyagarajan S.P., Mohanavalli B., Menon T.,
RA Froesner G.;
RT "PCR amplification, cloning and sequence determination of a hepatitis
RT E virus isolate from Madras, India.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X99441; CAA67804.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70960 MW; DAC3DF95F91689F8 CRC64;

```

```

Query Match 99.0%; Score 3374; DB 12; Length 660;
Best Local Similarity 99.1%; Pred. No. 1.8e-212;
Matches 654; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 MRPRPILLLLMLPMLPAPPQPSGRRRRSGSGGFGWDRVDSQFFAIPYIHPTN 60
Db 1 MRPRPILLLLMLPMLPAPPQPSGRRRRSGSGGFGWDRVDSQFFAIPYIHPTN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQARPAVASRRRPTTAGAAPLTAVPAHDTT 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQARPAASRRRPTTAGAAPLTAVPAHDTT 120

```

```

QY 121 PVPDVDSRGAILLRQYNLSTPLTSSVATGCTNLVLYAAPLSPLLPLQDGTNTHIMATEAS 180
Db 121 PVPDVDSRGAILLRQYNLSTPLTSSVATGCTNLVLYAAPLSPLLPLQDGTNTHIMATEAS 180
QY 181 NVAQYRVARATIRYRLPVNAVGGYAISISFWPQTITPTSDVMNSITSTDVRLVQPGI 240
Db 181 NVAQYRVARATIRYRLPVNAVGGYAISISFWPQTITPTSDVMNSITSTDVRLVQPGI 240
QY 241 ASSELVIPSRLHYRNQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGTALGLL 300
Db 241 ASSELVIPSRLHYRNQWRSVETSGVAEEATSGLVMLCIHGSPVNSYNTPTGTALGLL 300
QY 301 DFALELEFNLPNGTNTVRSYSSSTARHRLRGADGTAELTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFNLPNGTNTVRSYSSSTARHRLRGADGTAELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDLGESRVVIQDYDQHQEDRPTSPAPSRPFSVLRANDVLMSLTAAEY 480
Db 421 QDKGIAIPHDIIDLGESRVVIQDYDQHQEDRPTSPAPSRPFSVLRANDVLMSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLQGRPLSTIQOYKSTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLQGRPLSTIQOYKSTFFVLP 540
QY 541 LRKGLSFWAGTTKAGYPYNYNTASDQLLVENAGHRVAISYTTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWAGTTKAGYPYNYNTASDQLLVENAGHRVAISYTTTSLGAGPVSISAVAV 600
QY 601 LAPHSALEDTLDYPARAHTFDDFCPCPRPLGQCAFQSTVABLQRLKMKVGKTRREL 660
Db 601 LAPHSALEDTLDYPARAHTFDDFCPCPRPLGQCAFQSTVABLQRLKMKVGKTRREL 660

RESULT 8
Q81878
ID Q81878 PRELIMINARY; PRT; 660 AA.
AC Q81878;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Complete genome sequence.
OS Hepatitis E virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Uighi179;
RA Uchida T.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; D11093; BAA01867.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70937 MW; 5412313F1A2EF4A9 CRC64;

```

```

Query Match 98.9%; Score 3372; DB 12; Length 660;
Best Local Similarity 98.6%; Pred. No. 2.4e-212;
Matches 651; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 MRPRPILLLLMLPMLPAPPQPSGRRRRSGSGGFGWDRVDSQFFAIPYIHPTN 60
Db 1 MRPRPILLLLMLPMLPAPPQPSGRRRRSGSGGFGWDRVDSQFFAIPYIHPTN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQARPAVASRRRPTTAGAAPLTAVPAHDTT 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQARPAASRRRPTTAGAAPLTAVPAHDTT 120

```

```
QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVAAPLSPLLPLQDGTNTTHIMATEAS 180
Db 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVAAPLSPLLPLQDGTNTTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTITPTTSVDMNSITSDVRLVQPGI 240
Db 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTITPTTSVDMNSITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGTGALL 300
Db 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGTGALL 300
QY 301 DFALELEFRNLTPCNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFRNLTPCNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNADTLGGLPTELISSAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNADTLGGLPTELISSAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRWVIQDYDNOHQEODRPTSPAPSRPFSVLRANDVWLSLTAAEY 480
Db 421 QDKGIAIPHDIDLGESRWVIQDYDNOHQEODRPTSPAPSRPFSVLRANDVWLSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIOQYKTFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIOQYKTFVLP 540
QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAAHRAVISTVTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAAHRAVISTVTTSLGAGPVSISAVAV 600
QY 601 LAPHSAALLEDLTDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660
Db 601 LAPHSAALLEDLTDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660
```

RESULT 9

```
O91856 PRELIMINARY; PRT; 660 AA.
ID O91856
AC O91856;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=94-Egypt;
RX MEDLINE=99105430; PubMed=9890424;
RA Tsarev S.A., Binn L.N., Gomatos P.J., Arthur R.R., Monier M.K.,
RA van Cuyck-Gandre H., Longer C.F., Innis B.L.;
RT "Phylogenetic analysis of hepatitis E virus isolates from Egypt.";
RL J. Med. Virol. 57:68-74(1999).
DR EMBL; AF051352; AAC35744.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71026 MW; 16C560FA16941F2A CRC64;
```

Query Match 98.8%; Score 3367; DB 12; Length 660;
Best Local Similarity 98.8%; Pred. No. 5.1e-212;
Matches 652; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```
QY 1 MRPRPILLUMFLPMLPAPPQPSGRRRRGSGGPGWGRVDSQPPAIPYIHTN 60
Db 1 MRPRPILLUMFLPMLPAPPQPSGRRRRGSGGPGWGRVDSQPPAIPYIHTN 60
QY 61 PFAPDVTAAAGAPRVQRPARPLGSAMRDQAQRPVASSRRPTTAGAAPTAVAPADHTP 120
```

```
Db 61 PFAPDVTAAAGAPRVQRPARPLGSAMRDQAQRPVASSRRPTTAGAAPTAVAPADHTP 120
QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVAAPLSPLLPLQDGTNTTHIMATEAS 180
Db 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVAAPLSPLLPLQDGTNTTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTITPTTSVDMNSITSDVRLVQPGI 240
Db 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTITPTTSVDMNSITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGTGALL 300
Db 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTGTGALL 300
QY 301 DFALELEFRNLTPCNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFRNLTPCNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNADTLGGLPTELISSAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNADTLGGLPTELISSAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRWVIQDYDNOHQEODRPTSPAPSRPFSVLRANDVWLSLTAAEY 480
Db 421 QDKGIAIPHDIDLGESRWVIQDYDNOHQEODRPTSPAPSRPFSVLRANDVWLSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIOQYKTFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIOQYKTFVLP 540
QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAAHRAVISTVTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAAHRAVISTVTTSLGAGPVSISAVAV 600
QY 601 LAPHSAALLEDLTDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660
Db 601 LAPHSAALLEDLTDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660
```

RESULT 10

```
Q69411 PRELIMINARY; PRT; 660 AA.
ID Q69411
AC Q69411;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=hev037;
RX Donati M.C., Fagan E.A., Harrison T.J.;
RT "Sequence analysis of full length HEV clones derived directly from
RT human liver in fulminant hepatitis E.";
RL (In) Rizzetto M., Purcell R.H., Gerin J.L., Verme G (eds.);
RL VIRAL HEPATITIS AND LIVER DISEASE, pp.313-316,
RL Edizioni Minerva Medica, Torino (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=hev037;
RX Harrison T.J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X98292; CAA66937.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71024 MW; F37E3FEFF7A4EAD1 CRC64;
```

Query Match	98.8%;	Score 3367;	DB 12;	Length 660;
Best Local Similarity	98.5%;	Pred. No. 5.1e-212;		
Matches 650;	Conservative 4;	Mismatches 6;	Indels 0;	Gaps 0;
QY	1	MRPRPILLLLLMFLPMLPAPPQPSGRRRRSGSGGGFGWDRVDSQPPFAIPYIHPN	60	
Db	1	MRPRPILLLLLMFLPMLPAPPQPSGRRRRSGSGGGFGWDRVDSQPPFAIPYIHPN	60	
QY	61	PPAPDVTAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADTP	120	
Db	61	PPAPDVTAAGAGPRVROPAPPLGSAWRDQAPPAATSRRTTAGAAPTAVAPADTP	120	
QY	121	PVPDVDSGAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS	180	
Db	121	PVPDVDSGAILRRQYNLSTPLTSPVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS	180	
QY	181	NYAQYRVARATIRYRPLVNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIILVQPGI	240	
Db	181	NYAQYRVARATIRYRPLVNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIILVQPGI	240	
QY	241	ASELVI PSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGLPWSYNTPTGTALGLL	300	
Db	241	ASELVI PSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGLPWSYNTPTGTALGLL	300	
QY	301	DFALELEFRNLTPGNTNTRVSYSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG	360	
Db	301	DFALEFEFNLTGNTNTRVSYSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG	360	
QY	361	VEIGIRGIALTLFNLADTLGLLPTLTISSAGGQLFYSRPVVSANGEPTVKLYTSVENAQ	420	
Db	361	VEIGIRGIALTLFNLADTLGLLPTLTISSAGGQLFYSRPVVSANGEPTVKLYTSVENAQ	420	
QY	421	QDKGIAIPHIDILGESRVVIQDYDNOHQEODRTPSPAPSRPFSVLRANDVLSLTAAY	480	
Db	421	QDKGIAIPHIDILGESRVVIQDYDNOHQEODRTPSPAPSRPFSVLRANDVLSLTAAY	480	
QY	481	DOSTYGSSTGTPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRLSTIIQQYKTFVFLP	540	
Db	481	DOSTYGSSTGTPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRLSTIIQQYKTFVFLP	540	
QY	541	LRGKLSFWEAGTTKAGYPYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVVISAVAV	600	
Db	541	LRGKLSFWEAGTTKAGYPYNTTASDQLLIENAAAGHRVAISTYTTSLGAGPVVAISAVAV	600	
QY	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKTRREL	660	
Db	601	LAPHSALALLEDTMDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKTRREL	660	
RESULT 11				
Q9WLL4	PRELIMINARY; PRT; 660 AA.			
ID	Q9WLL4			
AC	Q9WLL4;			
DT	01-NOV-1999 (TremBLrel. 12, Created)			
DT	01-NOV-1999 (TremBLrel. 12, Last sequence update)			
DT	01-OCT-2003 (TremBLrel. 25, Last annotation update)			
DE	Capsid protein.			
OS	Hepatitis E virus.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage;			
OC	Hepatitis E-like viruses.			
OX	NCBI_TaxID=12461;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=TK15/92;			
RX	MEDLINE=99049628; PubMed=9833882;			
RA	Gouvea V., Snellings N., Popek M.J., Longer C.F., Innis B.L.;			
RT	"Hepatitis E virus: complete genome sequence and phylogenetic analysis			
RT	of a Nepali isolate.";			
RL	Virus Res. 57:21-26(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=TK15/92;			
RA	Gouvea V.;			

Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.				
EMBL; AF051830; AAC97188.1; -				
GO; GO:0005198; F:structural molecule activity; IEA.				
InterPro; IPR004261; SP2.				
InterPro; IPR008975; Viral_cap_coat.				
Pfam; PF03014; SP2; 1.				
SEQUENCE 660 AA; 70798 MW; A04C0185ACC085DB CRC64;				
Query Match	98.7%;	Score 3362;	DB 12;	Length 660;
Best Local Similarity	98.8%;	Pred. No. 1.1e-211;		
Matches 652;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;
QY	1	MRPRILLLLMFLPMLPAPPQPSGRRRRSGSGGGFGWDRVDSQPPFAIPYIHPN	60	
Db	1	MRPRILLLLMFLPMLPAPPQPSGRRRRSGSGGGFGWDRVDSQPPFAIPYIHPN	60	
QY	61	PPAPDVTAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADTP	120	
Db	61	PPAPDVTAAGAGPRVROPAPPLGSAWRDQAPPAATSRRRPTTAGAAPTAVAPADTP	120	
QY	121	PVPDVDSGAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS	180	
Db	121	PVPDVSGRGAIIICRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS	180	
QY	181	NYAQYRVARATIRYRPLVNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIILVQPGI	240	
Db	181	NYAQYRVARAAIRYRPLVNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIILVQPGI	240	
QY	241	ASELVI PSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSPVNSYNTPTGTALGLL	300	
Db	241	ASELVI PSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSPVNSYNTPTGTALGLL	300	
QY	301	DFALELEFRNLTPGNTNTRVSYSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG	360	
Db	301	DFALELEFRNLTPGNTNTRVSYSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG	360	
QY	361	VEIGIRGIALTLFNLADTLGLLPTLTISSAGGQLFYSRPVVSANGEPTVKLYTSVENAQ	420	
Db	361	VEIGIRGIALTLFNLADTLGLLPTLTISSAGGQLFYSRPVVSANGEPTVKLYTSVENAQ	420	
QY	421	QDKGIAIPHIDILGESRVVIQDYDNOHQEODRTPSPAPSRPFSVLRANDVLSLTAAY	480	
Db	421	QDKGIAIPHIDILGESRVVIQDYDNOHQEODRTPSPAPSRPFSVLRANDVLSLTAAY	480	
QY	481	DOSTYGSSTGTPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRLPLSTIIQQYKTFEVL	540	
Db	481	DOSTYGSSTGTPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRLPLSTIIQQYKTFEVL	540	
QY	541	LRGKLSFWEAGTTKAGYPYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVVISAVAV	600	
Db	541	LRGKLSFWEAGTTKAGYPYNTTASGQLLVENAAAGHRVAISTYTTSLGAGPVVISAVAV	600	
QY	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKTRREL	660	
Db	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKTRREL	660	
RESULT 12				
Q8JUT32	PRELIMINARY; PRT; 659 AA.			
ID	Q8JUT32			
AC	Q8JUT32;			
DT	01-OCT-2002 (TremBLrel. 22, Created)			
DT	01-OCT-2002 (TremBLrel. 22, Last sequence update)			
DT	01-OCT-2003 (TremBLrel. 25, Last annotation update)			
DE	Capsid protein.			
OS	Hepatitis E virus.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage;			
OC	Hepatitis E-like viruses.			
OX	NCBI_TaxID=12461;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22072932; Pubmed=12076829;			
RA	Jameel S., Zafrullah M., Chawla Y.K., Dillawari J.B.;			

RT "Reevaluation of a North India isolate of hepatitis E virus based on
RT the full-length genomic sequence obtained following long RT-PCR";
RL Virus Res. 86:53-58(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Jameel S., Zafrullah M., Chawla Y.K., Dilawari J.B.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF459438; AAM66330.1; -
DR EMBL; AF459438; AAM66330.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 659 AA; 70842 MW; ABCC177BBE6012C8 CRC64;

Query Match 98.18; Score 3342.5; DB 12; Length 659;
Best Local Similarity 98.3%; Pred. No. 2e-210;
Matches 649; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

QY 1 MRPRPILLLLMFLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPIHPTN 60
Db 1 MRPRPILLLLMFLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPIHPTN 59

QY 61 PFAPDVTAAGAGPRVQPARPLGSARDQAORPAVASRRRPTTAGAAPTAVAPADHTP 120
Db 60 PFAPDVTAAGAGPRVQPARPLGSARDQAORPAVASRRRPTTAGAAPTAVAPADHTP 119

QY 121 PVPDVSRGAILLRQYNLSPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 120 PVPDVSRGAILLRQYNLSPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 179

QY 181 NYAQYRVARATIRYRPLVNAVGYAISISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 240
Db 180 NYAQYRVARATIRYRPLVNAVGYAISISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 239

QY 241 ASELVTPSERLHYRNQWRSVETSGVAEEATSGLVMLCIHGSIVNSYTNTPYTGALGLL 300
Db 240 ASELVTPSERLHYRNQWRSVETSGVAEEATSGLVMLCIHGSIVNSYTNTPYTGALGLL 299

QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 300 DFALLEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 359

QY 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPYTKLYTSVENAQ 420
Db 360 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPYTKLYTSVENAQ 419

QY 421 QDKGIALPHDIDGESRVITQDYNQHEQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
Db 420 QDKGIALPHDIDGESRVITQDYNQHEQDRPTSPAPSRPFSVLRANDVWLSTAAEY 479

QY 481 DQSTYGSSTGPVVYSDSVTLNVNATGAQAVARSLDWTKVLDGRPLSTIQOYSKTFVLP 540
Db 480 DQSTYGSSTGPVVYSDSVTLNVNATGAQAVARSLDWTKVLDGRPLSTIQOYSKTFVLP 539

QY 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVV 600
Db 540 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVV 599

QY 601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPLGLQCAFQSTVAELQRLKMKVKTKREL 660
Db 600 LAPHSAALLEDTLDYPARAHTFDDFCPCRPLGLQCAFQSTVAELQRLKMKVKTKREL 659

RESULT 13
Q68985 PRELIMINARY; PRT; 660 AA.
AC Q68985;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF-2.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;

OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HYDERABAD;
RX MEDLINE=96087204; PubMed=8567900;
RA Panda S.K., Nanda S.K., Zafrullah M., Ansari I.H., Ozdener M.H.,
RA Jameel S.;
RT "An Indian strain of hepatitis E virus (HEV): cloning, sequence, and
RT expression of structural region and antibody responses in sera from
RT individuals from an area of high-level HEV endemicity";
RL J. Clin. Microbiol. 33:2653-2659(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Ansari I.H., Nanda S.K., Durgapal H., Jameel S., Panda S.K.;
RT "Translation analysis of complete HEV genome";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; U22532; AAA97366.1; -
DR EMBL; AF076239; AAC27936.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70936 MW; 948556F804382EB3 CRC64;

Query Match 97.9%; Score 3336; DB 12; Length 660;
Best Local Similarity 98.0%; Pred. No. 5.5e-210;
Matches 647; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMFLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPIHPTN 60
Db 1 MGPRPILLLLMFLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPIHPTN 60

QY 61 PFAPDVTAAGAGPRVQPARPLGSARDQAORPAVASRRRPTTAGAAPTAVAPADHTP 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSARDQAORPAVASRRRPTTAGAAPTAVAPADHTP 120

QY 121 PVPDVSRGAILLRQYNLSPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSRGAILLRQYNLSPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180

QY 181 NYAQYRVARATIRYRPLVNAVGYAISISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGYAISISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 240

QY 241 ASELVTPSERLHYRNQWRSVETSGVAEEATSGLVMLCIHGSIVNSYTNTPYTGALGLL 300
Db 241 ASELVTPSERLHYRNQWRSVETSGVAEEATSGLVMLCIHGSIVNSYTNTPYTGALGLL 300

QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360

QY 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPYTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPYTKLYTSVENAQ 420

QY 421 QDKGIALPHDIDGESRVITQDYNQHEQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
Db 421 QDKGIALPHDIDGESRVITQDYNQHEQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480

QY 481 DQSTYGSSTGPVVYSDSVTLNVNATGAQAVARSLDWTKVLDGRPLSTIQOYSKTFVLP 540
Db 481 DQSTYGSSTGPVVYSDSVTLNVNATGAQAVARSLDWTKVLDGRPLSTIQOYSKTFVLP 540

QY 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVV 600
Db 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVV 600

QY 601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPLGLQCAFQSTVAELQRLKMKVKTKREL 660
Db 601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPLGLQCAFQSTVAELQRLKMKVKTKREL 660

```
RESULT 14
Q8JUN6
ID Q8JUN6 PRELIMINARY; PRT; 674 AA.
AC Q8JUN6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ORF2 protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=JAK-Sai;
RA Mishiro S.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=JAK-Sai;
RX MEDLINE=21996248; PubMed=12001054;
RA Takahashi K., Kang J.-H., Ohnishi S., Hino K., Mishiro S.;
RT "Genetic Heterogeneity of Hepatitis E Virus Recovered from Japanese Patients with Acute Sporadic Hepatitis.";
RL J. Infect. Dis. 185:1342-1345(2002).
DR EMBL; AB074915; BAB96557.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 674 AA; 72591 MW; 5C196112F81A2265 CRC64;

Query Match 93.3%; Score 3179; DB 12; Length 674;
Best Local Similarity 92.1%; Pred. No. 1.1e-199;
Matches 607; Conservative 25; Mismatches 27; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGFGWDRVDSQPFAIPYIHPIN 60
DB 15 MRSRALLLFLLVFLPMLPAPPQPSGRRGRSGSGGFGWDRVDSQPFAIPYIHPIN 74
QY 61 PRAPDVTAAGAGPRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPTAVAPADHTP 120
DB 75 PRASDIPTAAGAGARPROPAPRLGSAWRDQSRPTASARRRPAAGAAPLTAVAPADTA 134
QY 121 PVPDVSRAILLRQYNLTSPLTSSVATGNTLVLYAAPLSPLLPDQDGTNTTHIMATEAS 180
DB 135 PVPDVSRAILLRQYNLTSPLTSSVATGNTLVLYAAPLSPLLPDQDGTNTTHIMATEAS 194
QY 181 NYAQVRVARATIRYRLVPLVNAVGGYAISSFPQTTTPTSDVMNSITSDVRLVQPGI 240
DB 195 NYAQVRVVRATIRYRLVPLVNAVGGYAISSFPQTTTPTSDVMNSITSDVRLVQPGI 254
QY 241 ASELVTPSELHYRNGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGLL 300
DB 255 ASELVTPSELHYRNGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGLL 314
QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEELTTAATRFMKDLYFTSTNG 360
DB 315 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEELTTAATRFMKDLYFTSTNG 374
QY 361 VGEIGRGIALTLFNLADTLGGIPLTELISAGQQLFYSRPVVSANGEPTVKLYTSVENAQ 420
DB 375 VGEVGRGIALTLFNLADTLGGIPLTELISAGQQLFYSRPVVSANGEPTVKLYTSVENAQ 434
QY 421 QDKGIAIPHDIDLGESRVVIQDYNQHEQDRTPSPAPSRPFVSLRANDVWLSLTAAY 480
DB 435 QDKGIAIPHDIDLGESRVVIQDYNQHEQDRTPSPAPSRPFVSLRANDVWLSLTAAY 494
QY 481 DQSTYSSSTGPPVVSVDVTLVNVATGAQAVARSIDWTKVTLDCRPLSTCOQSKTFVLP 540
DB 495 DQTYSSSTNPMVSDVTIVFNVATGAQGVSRSLDWSKVTLDCRPLTTIQQYSKTFVLP 554
QY 541 LRGLKSFWEAGTTKAGYPYNYNTTASDQLIVENAGHRVAISTYTTTSLGAGPVVISAVAV 600

Db 555 LRGLKSFWEAGTTKAGYPYNYNTTASDQLIVENAGHRVCISTYTTNIGSGPVVISAVGV 614
QY 601 LAPHSALALLEDTLDYPARAHTFDDFCPCRCPLGLOGCAFOSTVAELQRLKMKVGKTR 659
DB 615 LAPHSALAALEDVDYPARAHTFDDFCPCRCPLGLOGCAFOSTVAELQRLKMKVGKTR 673

RESULT 15
Q8JUM8
ID Q8JUM8 PRELIMINARY; PRT; 660 AA.
AC Q8JUM8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF2 protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=JKN-Sap;
RA Mishiro S.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=JKN-Sap;
RX MEDLINE=21996248; PubMed=12001054;
RA Takahashi K., Kang J.-H., Ohnishi S., Hino K., Mishiro S.;
RT "Genetic Heterogeneity of Hepatitis E Virus Recovered from Japanese Patients with Acute Sporadic Hepatitis.";
RL J. Infect. Dis. 185:1342-1345(2002).
DR EMBL; AB074918; BAB96562.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70980 MW; 19233C22F3E57FF8 CRC64;

Query Match 93.2%; Score 3176; DB 12; Length 660;
Best Local Similarity 92.1%; Pred. No. 1.6e-199;
Matches 607; Conservative 26; Mismatches 26; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGFGWDRVDSQPFAIPYIHPIN 60
DB 1 MRPAVLLLLFLLVLLPMLPAPPQPSGRRGRSGSGGFGWDRVDSQPFAIPYIHPIN 60
QY 61 PRAPDVTAAGAGPRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPTAVAPADHTP 120
DB 61 PRADVVVSQAGATRQPPRLGSAWRDQSRPSAARRRSPAPAGAAPLTAVSPADTA 120
QY 121 PVPDVSRAILLRQYNLTSPLTSSVATGNTLVLYAAPLSPLLPDQDGTNTTHIMATEAS 180
DB 121 PVPDVSRAILLRQYNLTSPLTSSVATGNTLVLYAAPLSPLLPDQDGTNTTHIMATEAS 180
QY 181 NYAQVRVARATIRYRLVPLVNAVGGYAISSFPQTTTPTSDVMNSITSDVRLVQPGI 240
DB 181 NYAQVRVVRATIRYRLVPLVNAVGGYAISSFPQTTTPTSDVMNSITSDVRLVQPGI 240
QY 241 ASELVTPSELHYRNGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGLL 300
DB 241 ASELVTPSELHYRNGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGLL 300
QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEELTTAATRFMKDLYFTSTNG 360
DB 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGIPLTELISAGQQLFYSRPVVSANGEPTVKLYTSVENAQ 420
DB 361 VGEVGRGIALTLFNLADTLGGIPLTELISAGQQLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYNQHEQDRTPSPAPSRPFVSLRANDVWLSLTAAY 480
```

Db	421	ODKGITPHDIDLGDSRVWIQDYDNQHEQDRPTFSPAPSRPFVLRANDVLMLSLTAAEY	480
Qy	481	DQSTYGSSTGPFVYVSDSVTLVNVATGAQAVARSLDWTYKVTLDGRPLSTIQOYSKTFVLP	540
Db	481	DQTTYGSSTNPMYVSDTVTLVNVATGAQAVARSLDWSKVTLDGRPLTTIQOYSKTFVLP	540
Qy	541	LRGKLSFEAGTTKAGYPYNYNTTASDOLLVENNAAGHRVAISTYTTSLGAGPVSISAVAV	600
Db	541	LRGKLSFEAGTTKAGYPYNYNTTASDQILIENAAAGHRVAISTYTTSLGAGPTSISAVGV	600
Qy	601	LAPHSALALLEDTLDYPARAHTFDDFCPECRPLGLOGCAFQSTVAELQRLKMKYKGTRE	659
Db	601	LAPHSALAVLEDTVDYPARAHTFDDFCPECTILGLOGCAFQSTIAELQRLKMKYKGTRE	659

Search completed: August 16, 2004, 13:35:04
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 13:33:52 ; Search time 19 seconds
(without alignments)
1793.323 Million cell updates/sec

Title: US-09-851-410A-8

Perfect score: 3408

Sequence: 1 MRPRPILLLLMLPMLPAP.....QSTVAELQRLKMKVKTREL 660

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 segs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A COMB.pap:*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap:*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap:*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pap:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3408	100.0	660	1	US-08-240-049B-15
2	3408	100.0	660	1	US-08-259-148A-19
3	3408	100.0	660	1	US-08-484-054-19
4	3408	100.0	660	2	US-07-876-941A-19
5	3408	100.0	660	3	US-08-478-507-8
6	3408	100.0	660	3	US-08-542-634-13
7	3408	100.0	660	3	US-09-128-275A-8
8	3408	100.0	660	3	US-08-477-292-13
9	3408	100.0	660	4	US-09-553-427-8
10	3408	100.0	660	4	US-09-462-606-50
11	3408	100.0	660	4	US-07-870-985A-19
12	3408	100.0	660	5	PCT-US95-13703-13
13	3399	99.7	660	4	US-09-172-699-2
14	3391	99.5	660	4	US-09-462-606-12
15	3382	99.2	660	3	US-08-840-316-2
16	3382	99.2	660	3	US-08-809-523-2
17	3382	99.2	660	3	US-08-471-971-2
18	3382	99.2	660	4	US-09-462-606-49
19	3382	99.2	660	4	US-09-462-606-56
20	3382	99.2	660	4	US-09-402-776-2
21	3382	99.2	660	5	PCT-US93-08849A-2
22	3382	99.2	660	4	US-09-462-606-51
23	3374	99.0	660	4	US-09-462-606-52
24	3372	98.9	660	4	US-09-462-606-53
25	3367	98.8	660	4	US-09-462-606-55
26	3367	98.8	660	4	US-09-462-606-55
27	3336	97.9	660	4	US-09-462-606-54

28 3221 94.5 660 3 US-08-542-634-14 Sequence 14, Appl
29 3221 94.5 660 5 PCT-US95-13703-14 Sequence 14, Appl
30 3215 94.3 660 4 US-09-462-606-48 Sequence 48, Appl
31 3194.5 93.7 659 1 US-08-240-049B-16 Sequence 16, Appl
32 3194.5 93.7 659 1 US-08-259-148A-20 Sequence 20, Appl
33 3194.5 93.7 659 1 US-08-484-054-20 Sequence 20, Appl
34 3194.5 93.7 659 2 US-07-876-941A-20 Sequence 20, Appl
35 3194.5 93.7 659 3 US-08-477-292-14 Sequence 14, Appl
36 3194.5 93.7 659 4 US-07-870-985A-20 Sequence 20, Appl
37 3164 84.8 660 4 US-09-462-606-57 Sequence 57, Appl
38 2865 84.1 561 4 US-09-172-699-20 Sequence 20, Appl
39 2821 82.8 552 4 US-09-172-699-16 Sequence 16, Appl
40 2813 82.5 549 3 US-08-542-634-15 Sequence 15, Appl
41 2813 82.5 549 3 US-08-477-292-15 Sequence 15, Appl
42 2813 82.5 549 5 PCT-US95-13703-15 Sequence 15, Appl
43 2769 81.2 540 3 US-08-542-634-25 Sequence 25, Appl
44 2769 81.2 540 5 PCT-US95-13703-25 Sequence 25, Appl
45 2694 79.0 525 3 US-08-542-634-27 Sequence 27, Appl

ALIGNMENTS

RESULT 1

US-08-240-049B-15
; Sequence 15, Application US/08240049B
; Patent No. 5686239
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Tam, Albert W.
; APPLICANT: Yarbough, Patricia O.
; TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240.049B
; FILING DATE: 09-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles K. Sholtz
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.35; G32P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
; INDIVIDUAL ISOLATE: ORF-2
US-08-240-049B-15

Query Match 100.0%; Score 3408; DB 1; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRILLLLMLPMLPAPPPQSGRRRRSGSGGFWGDRVDSQPPAIPYIHTN 60

Db 1 MRPRPILLLLMFLMPLPAPPQPSGRRGRSGSGGGFWGDRVDSQFPAIPIHPTN 60
QY 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPAHDT 120
Db 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPAHDT 120
QY 121 PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLVLAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLVLAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTPTTPTSDVMSITSDVRLVQPGI 240
Db 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTPTTPTSDVMSITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSRVETSGVAEEBATSGLVMLCIHGSVLNSYNTPTYGALGL 300
Db 241 ASELVIPSERLHYRNOGWSRVETSGVAEEBATSGLVMLCIHGSVLNSYNTPTYGALGL 300
QY 301 DPALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DPALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISAGAGQLFYRPPVWSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTELISAGAGQLFYRPPVWSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDGESRVVIQDYNQHEQDRPTSPAPSPFVSLRANDVWLSLTAAEY 480
Db 421 QDKGIAIPHDIIDGESRVVIQDYNQHEQDRPTSPAPSPFVSLRANDVWLSLTAAEY 480
QY 481 DQSTYGSSTGVPVYSDSVTLVNVATGAQAVARSLDWTKVLDRPLSTIOQYKTFVLP 540
Db 481 DQSTYGSSTGVPVYSDSVTLVNVATGAQAVARSLDWTKVLDRPLSTIOQYKTFVLP 540
QY 541 LRKGLSFEWAGTTKAGYPYNNVTASDQLLVENAGHRVAISTVTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFEWAGTTKAGYPYNNVTASDQLLVENAGHRVAISTVTTSLGAGPVSISAVAV 600
QY 601 LAPHASALALLEDTLDYPARAHTFDDFCPCRPLGQCAFQSTVAELQRLKMKVGKTREL 660
Db 601 LAPHASALALLEDTLDYPARAHTFDDFCPCRPLGQCAFQSTVAELQRLKMKVGKTREL 660

RESULT 2

US-08-259-148A-19
; Sequence 19, Application US/08259148A
; Patent No. 5741490

GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Two, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; APPLICANT: Varbough, Patrice D.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/259,148A
; FILING DATE: 13-JUN-1994

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA: US 505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA: US 420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA: US 367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA: US 336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA: US 208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
; US-08-259-148A-19

Query Match 100.0%; Score 3408; DB 1; Length 660;

Best Local Similarity 100.0%; Pred. No. 3.7e-294;

Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMFLMPLPAPPQPSGRRGRSGSGGGFWGDRVDSQFPAIPIHPTN 60
Db 1 MRPRPILLLLMFLMPLPAPPQPSGRRGRSGSGGGFWGDRVDSQFPAIPIHPTN 60
QY 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPAHDT 120
Db 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPAHDT 120
QY 121 PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLVLAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLVLAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTPTTPTSDVMSITSDVRLVQPGI 240
Db 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTPTTPTSDVMSITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSRVETSGVAEEBATSGLVMLCIHGSVLNSYNTPTYGALGL 300
Db 241 ASELVIPSERLHYRNOGWSRVETSGVAEEBATSGLVMLCIHGSVLNSYNTPTYGALGL 300
QY 301 DPALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DPALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISAGAGQLFYRPPVWSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTELISAGAGQLFYRPPVWSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDGESRVVIQDYNQHEQDRPTSPAPSPFVSLRANDVWLSLTAAEY 480
Db 421 QDKGIAIPHDIIDGESRVVIQDYNQHEQDRPTSPAPSPFVSLRANDVWLSLTAAEY 480

```
QY 481 DQSTYGSSTGPPVYVSDSVTLNVNATCAQAVARSLDWTKVTLGDRPLSTIQQYSKTFVLP 540
Db 481 DQSTYGSSTGPPVYVSDSVTLNVNATCAQAVARSLDWTKVTLGDRPLSTIQQYSKTFVLP 540
QY 541 LRKLSFWEAGTTKAGYPYNNYNTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKLSFWEAGTTKAGYPYNNYNTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDTLDYPARAHTFDDFCPECRPLGLQGCAFSQSTVAELQRLKMKVGKTR 660
Db 601 LAPHSALELLEDTLDYPARAHTFDDFCPECRPLGLQGCAFSQSTVAELQRLKMKVGKTR 660

RESULT 3
US-08-484-054-19
; Sequence 19, Application US/08484054
; Patent No. 5770689
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,054
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 870,985
; FILING DATE: 20-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
US-08-484-054-19

Query Match 100.0%; Score 3408; DB 1; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294; Indels 0; Gaps 0;
Matches 660; Conservative 0; Mismatches 0;

QY 1 MRPRPILLLLMLFPLPAPPQPSGRRRRSGSGGFGWDRVDSQPPAIPYIHPIN 60
Db 1 MRPRPILLLLMLFPLPAPPQPSGRRRRSGSGGFGWDRVDSQPPAIPYIHPIN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSARWDOQRPVAVASRRRPTTAGAPLTAVAPAHDP 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSARWDOQRPVAVASRRRPTTAGAPLTAVAPAHDP 120
QY 121 PVPDVDSRGAILRROYNLSTSLTSSVATGTLNLYAAPLSPLLPLOGDTNTHIMATEAS 180
Db 121 PVPDVDSRGAILRROYNLSTSLTSSVATGTLNLYAAPLSPLLPLOGDTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVPNAVGGYAISISFWPQTTTPTSVDMSISITSDVRILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVPNAVGGYAISISFWPQTTTPTSVDMSISITSDVRILVQPGI 240
QY 241 ASELVIPSERLHYRNQGWRSVETSGVAEEATSGLMCLIHGSLVNSYNTPTYGALGILL 300
Db 241 ASELVIPSERLHYRNQGWRSVETSGVAEEATSGLMCLIHGSLVNSYNTPTYGALGILL 300
QY 301 DFALELEFNLTPGNTNTRVSRYSSTASHRLREGADGTAELTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFNLTPGNTNTRVSRYSSTASHRLREGADGTAELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLGLLPGTSSAGQQLFYSPRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLGLLPGTSSAGQQLFYSPRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAI PHDIDLGESRVVIQDYDQHEQDRTPSPAPSRPFSVLRLANDVLMISLTAAY 480
Db 421 QDKGIAI PHDIDLGESRVVIQDYDQHEQDRTPSPAPSRPFSVLRLANDVLMISLTAAY 480
QY 481 DQSTYGSSTGPPVYVSDSVTLNVNATCAQAVARSLDWTKVTLGDRPLSTIQQYSKTFVLP 540
Db 481 DQSTYGSSTGPPVYVSDSVTLNVNATCAQAVARSLDWTKVTLGDRPLSTIQQYSKTFVLP 540
QY 541 LRKLSFWEAGTTKAGYPYNNYNTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKLSFWEAGTTKAGYPYNNYNTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDTLDYPARAHTFDDFCPECRPLGLQGCAFSQSTVAELQRLKMKVGKTR 660
Db 601 LAPHSALELLEDTLDYPARAHTFDDFCPECRPLGLQGCAFSQSTVAELQRLKMKVGKTR 660

RESULT 4
US-07-876-941A-19
; Sequence 19, Application US/07876941A
; Patent No. 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
```



```
/ FILING DATE: 17-JUN-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Sholtz, Charles K.
/ REGISTRATION NUMBER: 38,615
/ REFERENCE/DOCKET NUMBER: 4600-0183.22
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 324-0880
/ TELEFAX: (650) 324-0960
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-478-507-8

Query Match      100.0%; Score 3408; DB 3; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPYIHPN 60
Db 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPYIHPN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADTP 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADTP 120
QY 121 PVPDVSRCAILRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSRCAILRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYRPLVNAVGVYAIISFPQTTTTPTSDVMSITSDVRILVOPGI 240
Db 181 NYAQVRVARATIRYRPLVNAVGVYAIISFPQTTTTPTSDVMSITSDVRILVOPGI 240
QY 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSLSVNSYNTPTGTALGLL 300
Db 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSLSVNSYNTPTGTALGLL 300
QY 301 DFALELEFNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGEPVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGEPVKLYTSVENAQ 420
QY 421 ODKGIAIPHDIDLGESRVVIQDYDNQHQDRPTSPAPSRPFVSLRANDVLMLSLTAAY 480
Db 421 ODKGIAIPHDIDLGESRVVIQDYDNQHQDRPTSPAPSRPFVSLRANDVLMLSLTAAY 480
```

RESULT 6

```
US-08-542-634-13
; Sequence 13, Application US/08542634
; Patent No. 6214970
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; APPLICANT: McAtee, C. Patrick
; APPLICANT: Yarborough, Patrice O.
```

```
/ APPLICANT: Zhang, Yifan
/ TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dehlinger & Associates
/ STREET: 350 Cambridge Ave., Suite 250
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94306
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/542,634
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0293.30
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
/ INDIVIDUAL ISOLATE: ORF-2
US-08-542-634-13
```

```
Query Match      100.0%; Score 3408; DB 3; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPYIHPN 60
Db 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPYIHPN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADTP 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADTP 120
QY 121 PVPDVSRCAILRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSRCAILRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYRPLVNAVGVYAIISFPQTTTTPTSDVMSITSDVRILVOPGI 240
Db 181 NYAQVRVARATIRYRPLVNAVGVYAIISFPQTTTTPTSDVMSITSDVRILVOPGI 240
QY 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSLSVNSYNTPTGTALGLL 300
Db 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSLSVNSYNTPTGTALGLL 300
QY 301 DFALELEFNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGEPVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGEPVKLYTSVENAQ 420
QY 421 ODKGIAIPHDIDLGESRVVIQDYDNQHQDRPTSPAPSRPFVSLRANDVLMLSLTAAY 480
Db 421 ODKGIAIPHDIDLGESRVVIQDYDNQHQDRPTSPAPSRPFVSLRANDVLMLSLTAAY 480
```

```
QY 481 DQSTYGSSTGPGVYVSDSVTLVNVATGAQAVARSLDWTKVTLDRPLSTIQOYKSTFFVLP 540
Db 481 DQSTYGSSTGPGVYVSDSVTLVNVATGAQAVARSLDWTKVTLDRPLSTIQOYKSTFFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHRAVIASTVTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHRAVIASTVTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDTLDYPARAHTFDDFCPECRPLGLQCAFGSTVAELQRLKMKVGTREL 660
Db 601 LAPHSALELLEDTLDYPARAHTFDDFCPECRPLGLQCAFGSTVAELQRLKMKVGTREL 660

RESULT 7
US-09-128-275A-8
; Sequence 8, Application US/09128275A
; Patent No. 6229005
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6229005-A/No. 6229005-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,275A
; FILING DATE: 03-AUG-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-128-275A-8

Query Match 100.0%; Score 3408; DB 3; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMFLPMLPAPPQPSGRRRGRSGGGGFGWGDVDSQFPALPYIHPN 60
Db 1 MRPRPILLLLMFLPMLPAPPQPSGRRRGRSGGGGFGWGDVDSQFPALPYIHPN 60

QY 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAQPAVASRRRPTTAGAAPTATAPAHDT 120
Db 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAQPAVASRRRPTTAGAAPTATAPAHDT 120

QY 121 PVPDVSRGAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSRGAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180

QY 181 NYAQYRVARATIRYRPLVPNAVGGYAIISIFWPTTTTPTSDVMSITSDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRPLVPNAVGGYAIISIFWPTTTTPTSDVMSITSDVRLVQPGI 240

QY 241 ASELVIPSERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSIVNSYNTPTYGALGL 300
Db 241 ASELVIPSERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSIVNSYNTPTYGALGL 300

QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATFRMKDLYFTSTNG 360
Db 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATFRMKDLYFTSTNG 360

QY 361 VGEIGRGIALTLFNADTLGLPTELISSAGGOLFYSRVPVVSANGEPVTKLYTSVNAQ 420
Db 361 VGEIGRGIALTLFNADTLGLPTELISSAGGOLFYSRVPVVSANGEPVTKLYTSVNAQ 420

QY 421 QDKGIAIPHDIIDIGESRVVIQDYDNOHQEODRPTSPAPSRPFVLRANDVILSLTAAEY 480
Db 421 QDKGIAIPHDIIDIGESRVVIQDYDNOHQEODRPTSPAPSRPFVLRANDVILSLTAAEY 480

QY 481 DQSTYGSSTGPGVYVSDSVTLVNVATGAQAVARSLDWTKVTLDRPLSTIQOYKSTFFVLP 540
Db 481 DQSTYGSSTGPGVYVSDSVTLVNVATGAQAVARSLDWTKVTLDRPLSTIQOYKSTFFVLP 540

QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHRAVIASTVTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHRAVIASTVTTSLGAGPVSISAVAV 600

QY 601 LAPHSALELLEDTLDYPARAHTFDDFCPECRPLGLQCAFGSTVAELQRLKMKVGTREL 660
Db 601 LAPHSALELLEDTLDYPARAHTFDDFCPECRPLGLQCAFGSTVAELQRLKMKVGTREL 660

RESULT 8
US-08-477-292-13
; Sequence 13, Application US/08477292
; Patent No. 6291641
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas
; APPLICANT: McAtee, Patrick
; APPLICANT: Yarbough, Patrice
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cathleen M. Desjardins, M.D.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: US
; ZIP: 94063
```

```
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/477,292
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/327,952
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Brookes, Allan A
/ REGISTRATION NUMBER: 36,373
/ REFERENCE/DOCKET NUMBER: G32P5
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415)369-9500
/ TELEFAX: (415)368-0709
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
/ INDIVIDUAL ISOLATE: ORF-2
/
US-08-477-292-13
Query Match 100.0%; Score 3408; DB 3; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPAIPYIHPN 60
Db 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPAIPYIHPN 60
QY 61 PPAPDVTAAAGAPRVOPARPLGSAWRDQORPAVASRRRTTAGAAPLTAVAPAHDP 120
Db 61 PPAPDVTAAAGAPRVOPARPLGSAWRDQORPAVASRRRTTAGAAPLTAVAPAHDP 120
QY 121 PVPDVSRRGAILRRQYNLSPLTSVATGTLNLYAAPLSPLLPLODGTNTHIMATAS 180
Db 121 PVPDVSRRGAILRRQYNLSPLTSVATGTLNLYAAPLSPLLPLODGTNTHIMATAS 180
QY 181 NYAQYRVARATIRYRPLVPNAVGGYAISSFWPQTTTTPTSDVMSITSTDVRIIVQPGI 240
Db 181 NYAQYRVARATIRYRPLVPNAVGGYAISSFWPQTTTTPTSDVMSITSTDVRIIVQPGI 240
QY 241 ASELVIPSERLHYRNOGRSVETSGVABEETSGLVMLCIHGSLVNSYNTPTGALGL 300
Db 241 ASELVIPSERLHYRNOGRSVETSGVABEETSGLVMLCIHGSLVNSYNTPTGALGL 300
QY 301 DFALEFRNLTPGNTNTRVRSYSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALEFRNLTPGNTNTRVRSYSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLSLPTELISSAGGQLFYGRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLSLPTELISSAGGQLFYGRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGTAIPHDDILGSRVVIQYDQNOHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
Db 421 QDKGTAIPHDDILGSRVVIQYDQNOHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
QY 481 DQSTYGSSTGTPVYVSDSVTLNVAAGAVARSLDWTVTLDGRPLSTIQYKSTFFVLP 540
Db 481 DQSTYGSSTGTPVYVSDSVTLNVAAGAVARSLDWTVTLDGRPLSTIQYKSTFFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNNYNTTASDQLLVENAAAGHRVAISTYTTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNNYNTTASDQLLVENAAAGHRVAISTYTTTSLGAGPVSISAVAV 600
```

RESULT 9

```
US-09-553-427-8
; Sequence 8, Application US/09553427
; Patent No. 6379891
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbrough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; HEPATITIS VIRAL AGENT
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/553,427
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,507
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0183.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-553-427-8
```

```
Db 541 LRKLSFWEAGTTKAGYPYNNYNTTASDQLLVENAAAGHRVAISTYTTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDTLDYPARAHTFDDFCPECRPLGLQGCAFOSTVAELQRLKMKVGTREL 660
Db 601 LAPHSALELLEDTLDYPARAHTFDDFCPECRPLGLQGCAFOSTVAELQRLKMKVGTREL 660
```

Query Match 100.0%; Score 3408; DB 4; Length 660;
 Best Local Similarity 100.0%; Pred. No. 3.7e-294;
 Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMPLPAPPQPSGRGRGRSGGGGFWGDRVDSQPFALPYIHTPN 60
 DB 1 MRPRPILLLLMPLPAPPQPSGRGRGRSGGGGFWGDRVDSQPFALPYIHTPN 60

QY 61 PFAPDVTAAGAPRVQPARPLGSAWRDQAQPAVASRRRPTTAGAAPLTAAPAHDT 120
 DB 61 PFAPDVTAAGAPRVQPARPLGSAWRDQAQPAVASRRRPTTAGAAPLTAAPAHDT 120

QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLPLQDGNTHIMATEAS 180
 DB 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLPLQDGNTHIMATEAS 180

QY 181 NYAQYVARATIRYRPLVNAVGYAISISFWPQTPTTPTSDVMNSITSDVRLVQPGI 240
 DB 181 NYAQYVARATIRYRPLVNAVGYAISISFWPQTPTTPTSDVMNSITSDVRLVQPGI 240

QY 241 ASELVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSIVNSYNTPTGTALGLL 300
 DB 241 ASELVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSIVNSYNTPTGTALGLL 300

QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
 DB 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360

QY 361 VGEIGRGIALTLFNADTLGGLPTELISSAGQOLFYSRPVWSANGEPTVKLYTSVENAQ 420
 DB 361 VGEIGRGIALTLFNADTLGGLPTELISSAGQOLFYSRPVWSANGEPTVKLYTSVENAQ 420

QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVLMISLTAAY 480
 DB 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVLMISLTAAY 480

QY 481 DQSTYGSSTGPFVVSVDVTLNVNATGAQAVARSLDWTKVLDGRPLSTIOQYKTFVLP 540
 DB 481 DQSTYGSSTGPFVVSVDVTLNVNATGAQAVARSLDWTKVLDGRPLSTIOQYKTFVLP 540

QY 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAHRAVIAISTVTTSLGAGPVSISAVAV 600
 DB 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAHRAVIAISTVTTSLGAGPVSISAVAV 600

QY 601 LAPHSALELTDLDYPARAHTFDDFCPCRPLGQCAFQSTVAELQRLKMKVGTREL 660
 DB 601 LAPHSALELTDLDYPARAHTFDDFCPCRPLGQCAFQSTVAELQRLKMKVGTREL 660

RESULT 10

US-09-462-606-50
 ; Sequence 50, Application US/09462606
 ; Patent No. 6432408
 ; GENERAL INFORMATION:
 ; APPLICANT: MENG, XIANG-JIN
 ; APPLICANT: Emerson, Suzanne U.
 ; APPLICANT: Purcell, Robert H.
 ; FILE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
 ; FILE REFERENCE: 20264267US1
 ; CURRENT APPLICATION NUMBER: US/09/462,606
 ; CURRENT FILING DATE: 2000-06-12
 ; PRIOR APPLICATION NUMBER: US 60/053069
 ; PRIOR FILING DATE: 1997-07-18
 ; PRIOR APPLICATION NUMBER: PCT/US98/14665
 ; PRIOR FILING DATE: 1998-07-17
 ; NUMBER OF SEQ ID NOS: 65
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 50
 ; LENGTH: 660
 ; TYPE: PRT
 ; ORGANISM: Hepatitis E virus
 ; US-09-462-606-50

RESULT 11

US-07-870-985A-19
 ; Sequence 19, Application US/07870985A
 ; Patent No. 6455492
 ; GENERAL INFORMATION:
 ; APPLICANT: Reyes, Gregory R.
 ; APPLICANT: Bradley, Daniel W.
 ; APPLICANT: Twu, Ji-Shin
 ; APPLICANT: Purdy, Michael A.
 ; APPLICANT: Tam, Albert W.
 ; APPLICANT: Krawczynski, Krzysztof Z.
 ; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dehlinger & Associates
 ; STREET: 350 Cambridge Avenue, Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible


```
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/870,985A
/ FILING DATE: 20-APRIL-1992
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 822,335
/ FILING DATE: 17-JAN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 505,888
/ FILING DATE: 05-APRIL-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 420,921
/ FILING DATE: 13-OCTOBER-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 367,486
/ FILING DATE: 16-JUNE-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 336,672
/ FILING DATE: 11-APRIL-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 208,997
/ FILING DATE: 17-JUNE-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0093.30
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 19:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
/ US-07-870-985A-19

Query Match 100.0%; Score 3408; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRILLLLMLPAPPQSGRRRRGSGGFGWDRVDSQPPFAIPYIHPTN 60
Db 1 MRPRILLLLMLPAPPQSGRRRRGSGGFGWDRVDSQPPFAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQORPAVASRRRTTAGAAPLTAAPAHTDP 120
Db 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQORPAVASRRRTTAGAAPLTAAPAHTDP 120
QY 121 PVPDVSRRGAILRRQYNLSPTLSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEAS 180
Db 121 PVPDVSRRGAILRRQYNLSPTLSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRPLVPNAVGGVAISIFWPQTTPTSVDMNSITSDVRILVQPGI 240
Db 181 NYAQVRVARATIRPLVPNAVGGVAISIFWPQTTPTSVDMNSITSDVRILVQPGI 240
QY 241 ASELVIPSRLHYRNOGRSVETSGVAEEATSGVLMLCIHGSLVNSYNTPTGALGL 300
Db 241 ASELVIPSRLHYRNOGRSVETSGVAEEATSGVLMLCIHGSLVNSYNTPTGALGL 300
QY 301 DFALEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEELTTAATRFMKDLYFTSTNG 360
Db 301 DFALEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLGLPTLTISSAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLGLPTLTISSAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420

/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/870,985A
/ FILING DATE: 20-APRIL-1992
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 822,335
/ FILING DATE: 17-JAN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 505,888
/ FILING DATE: 05-APRIL-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 420,921
/ FILING DATE: 13-OCTOBER-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 367,486
/ FILING DATE: 16-JUNE-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 336,672
/ FILING DATE: 11-APRIL-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 208,997
/ FILING DATE: 17-JUNE-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0093.30
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 19:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/13703
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0293.41
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
/ INDIVIDUAL ISOLATE: ORF-2
/ PCT-US95-13703-13

Query Match 100.0%; Score 3408; DB 5; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRILLLLMLPAPPQSGRRRRGSGGFGWDRVDSQPPFAIPYIHPTN 60
Db 1 MRPRILLLLMLPAPPQSGRRRRGSGGFGWDRVDSQPPFAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQORPAVASRRRTTAGAAPLTAAPAHTDP 120
Db 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQORPAVASRRRTTAGAAPLTAAPAHTDP 120

/ RESULT 12
/ PCT-US95-13703-13
/ Sequence 13, Application PC/TUS9513703
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
/ TITLE OF INVENTION: USES THEREFOR
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dehlinger & Associates
/ STREET: P.O. Box 60850
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94306-0850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/13703
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0293.41
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
/ INDIVIDUAL ISOLATE: ORF-2
/ PCT-US95-13703-13

Query Match 100.0%; Score 3408; DB 5; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRILLLLMLPAPPQSGRRRRGSGGFGWDRVDSQPPFAIPYIHPTN 60
Db 1 MRPRILLLLMLPAPPQSGRRRRGSGGFGWDRVDSQPPFAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQORPAVASRRRTTAGAAPLTAAPAHTDP 120
Db 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQORPAVASRRRTTAGAAPLTAAPAHTDP 120
```

QY 121 PVPDVSRAILRRQYNLSPTSSVATGTLNLVYAAPLSPLPLQDGTNTHIMATEAS 180
 Db 121 PVPDVSRAILRRQYNLSPTSSVATGTLNLVYAAPLSPLPLQDGTNTHIMATEAS 180
 QY 181 NYAQYVARATIRYRPLVNAVGGYVAISISFMPQTTTTPTSDVMSITSDVRLVQPGI 240
 Db 181 NYAQYVARATIRYRPLVNAVGGYVAISISFMPQTTTTPTSDVMSITSDVRLVQPGI 240
 QY 241 ASLVTPSRLHYRNOGWSVETSGVAEBEATSGLVMLCIHGLSVNSYNTPTGTALGLL 300
 Db 241 ASLVTPSRLHYRNOGWSVETSGVAEBEATSGLVMLCIHGLSVNSYNTPTGTALGLL 300
 QY 301 DPALELEFNLTGNTNTRVSRYSSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
 Db 301 DPALELEFNLTGNTNTRVSRYSSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
 QY 361 VGEIGRGIALTLFNLADTLTLLGGLPTELISSAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
 Db 361 VGEIGRGIALTLFNLADTLTLLGGLPTELISSAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
 QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLANDVWLSTAAEY 480
 Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLANDVWLSTAAEY 480
 QY 481 DQSTYSSGTPVVSVDVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQOYSKTFVLP 540
 Db 481 DQSTYSSGTPVVSVDVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQOYSKTFVLP 540
 QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRAVIASTYTTSLGAGPVSISAVAV 600
 Db 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRAVIASTYTTSLGAGPVSISAVAV 600
 QY 601 LAPHSALELLEDTLPARAHTEDDFCPCRLGLOGCAFOSTVAELQRLKMKVGTREL 660
 Db 601 LAPHSALELLEDTLPARAHTEDDFCPCRLGLOGCAFOSTVAELQRLKMKVGTREL 660

RESULT 13

US-09-172-699-2
 ; Sequence 2, Application US/09172699A
 ; Patent No. 6514690
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, David A.
 ; APPLICANT: Locarnini, Stephen A.
 ; APPLICANT: Toressi, Joseph
 ; APPLICANT: Hui, Zhuang
 ; APPLICANT: Li, Fan
 ; TITLE OF INVENTION: IMMUNOREACTIVE ANTIGENS OF HEPATITIS E VIRUS
 ; FILE REFERENCE: Davies Col. Cave
 ; CURRENT APPLICATION NUMBER: US/09/172,699A
 ; CURRENT FILING DATE: 1998-10-14
 ; EARLIER APPLICATION NUMBER: 08/617,927
 ; EARLIER FILING DATE: 1996-06-20
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 660
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-172-699-2

Query Match 99.78; Score 3399; DB 4; Length 660;
 Best Local Similarity 99.78; Pred. No. 2.3e-293;
 Matches 658; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MRPRPILLLLMLFPLMPPPPQSGRRGRSGGGGFWGDRVDSQFAIPIYHPTN 60
 Db 1 MRPRPILLLLMLFPLMPPPPQSGRRGRSGGGGFWGDRVDSQFAIPIYHPTN 60
 QY 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQRPVAVASRRRPTTAGAAPTAVAPADHTP 120
 Db 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQRPVAVASRRRPTTAGAAPTAVAPADHTP 120

QY 121 PVPDVSRAILRRQYNLSPTSSVATGTLNLVYAAPLSPLPLQDGTNTHIMATEAS 180
 Db 121 PVPDVSRAILRRQYNLSPTSSVATGTLNLVYAAPLSPLPLQDGTNTHIMATEAS 180
 QY 181 NYAQYVARATIRYRPLVNAVGGYVAISISFMPQTTTTPTSDVMSITSDVRLVQPGI 240
 Db 181 NYAQYVARATIRYRPLVNAVGGYVAISISFMPQTTTTPTSDVMSITSDVRLVQPGI 240
 QY 241 ASLVTPSRLHYRNOGWSVETSGVAEBEATSGLVMLCIHGLSVNSYNTPTGTALGLL 300
 Db 241 ASLVTPSRLHYRNOGWSVETSGVAEBEATSGLVMLCIHGLSVNSYNTPTGTALGLL 300
 QY 301 DPALELEFNLTGNTNTRVSRYSSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
 Db 301 DPALELEFNLTGNTNTRVSRYSSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
 QY 361 VGEIGRGIALTLFNLADTLTLLGGLPTELISSAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
 Db 361 VGEIGRGIALTLFNLADTLTLLGGLPTELISSAGGOLFYSRPVVSANGEPTVKLYTSVENAQ 420
 QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLANDVWLSTAAEY 480
 Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLANDVWLSTAAEY 480
 QY 481 DQSTYSSGTPVVSVDVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQOYSKTFVLP 540
 Db 481 DQSTYSSGTPVVSVDVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQOYSKTFVLP 540
 QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRAVIASTYTTSLGAGPVSISAVAV 600
 Db 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRAVIASTYTTSLGAGPVSISAVAV 600
 QY 601 LAPHSALELLEDTLPARAHTEDDFCPCRLGLOGCAFOSTVAELQRLKMKVGTREL 660
 Db 601 LAPHSALELLEDTLPARAHTEDDFCPCRLGLOGCAFOSTVAELQRLKMKVGTREL 660

RESULT 14

US-09-462-606-12
 ; Sequence 12, Application US/09462606
 ; Patent No. 6432408
 ; GENERAL INFORMATION:
 ; APPLICANT: MENG, XIANG-JIN
 ; APPLICANT: Emerson, Suzanne U.
 ; APPLICANT: Purcell, Robert H.
 ; TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
 ; FILE REFERENCE: 20264267US1
 ; CURRENT APPLICATION NUMBER: US/09/462,606
 ; CURRENT FILING DATE: 2000-06-12
 ; PRIOR APPLICATION NUMBER: US 60/053069
 ; PRIOR FILING DATE: 1997-07-18
 ; PRIOR APPLICATION NUMBER: PCT/US98/14665
 ; PRIOR FILING DATE: 1998-07-17
 ; NUMBER OF SEQ ID NOS: 65
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 12
 ; LENGTH: 660
 ; TYPE: PRT
 ; ORGANISM: Hepatitis E virus
 US-09-462-606-12

Query Match 99.5%; Score 3391; DB 4; Length 660;
 Best Local Similarity 99.4%; Pred. No. 1.2e-292;
 Matches 656; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MRPRPILLLLMLFPLMPPPPQSGRRGRSGGGGFWGDRVDSQFAIPIYHPTN 60
 Db 1 MRPRPILLLLMLFPLMPPPPQSGRRGRSGGGGFWGDRVDSQFAIPIYHPTN 60
 QY 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQRPVAVASRRRPTTAGAAPTAVAPADHTP 120
 Db 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQRPVAVASRRRPTTAGAAPTAVAPADHTP 120

```
QY 121 PVPDVDSRGAILRRQVNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVDSRGAILRRQVNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYRPLVNPVNAVGGVAISISFWPQTTPPTTPTSDVMSITSTDVRLVQPGI 240
Db 181 NYAQVRVARATIRYRPLVNPVNAVGGVAISISFWPQTTPPTTPTSDVMSITSTDVRLVQPGI 240
QY 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVLNSYNTNTPTGTALGLL 300
Db 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVLNSYNTNTPTGTALGLL 300
QY 301 DFALBEPFNLTGNTNTRVSYSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
Db 301 DFALBEPFNLTGNTNTRVSYSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTLISAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTLISAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGTAIPHDIIDLGESRVVIQDYDQHEQDRPTSPAPSRPFSVLRANDVWLSLTAABY 480
Db 421 QDKGTAIPHDIIDLGESRVVIQDYDQHEQDRPTSPAPSRPFSVLRANDVWLSLTAABY 480
QY 481 LRGKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 481 LRGKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHASALALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGTREL 660
Db 601 LAPHASALALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGTREL 660
```

RESULT 15

US-08-840-316-2

; Sequence 2, Application US/08840316

; Patent No. 6054567

; GENERAL INFORMATION:

; APPLICANT: Emerson, Suzanne U., Purcell, Robert H.,

; APPLICANT: Tearev, Sergei. A., and Robinson, Robin A.

; TITLE OF INVENTION: Recombinant Proteins Of

; TITLE OF INVENTION: A Pakistani Strain Of Hepatitis E And Their

; TITLE OF INVENTION: Use In Diagnostic Methods And Vaccines

; NUMBER OF SEQUENCES: 111

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/840,316

; FILING DATE: 11-APR-1997

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Richard W. Bork

; REGISTRATION NUMBER: 36,459

; REFERENCE/DOCKET NUMBER: 2026-4255

; TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
US-08-840-316-2
```

Query Match 99.2%; Score 3382; DB 3; Length 660;

Best Local Similarity 99.2%; Pred. No. 7.6e-292;

Matches 655; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 MRPRLTLLLLMFLPMLPAPPQPSGRRGRSGSGGGFWGDRVDSOPEALPYIHPNT 60
Db 1 MRPRLTLLLLMFLPMLPAPPQPSGRRGRSGSGGGFWGDRVDSOPEALPYIHPNT 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSARWDAQORPAVASRRRPTTAGAAPLTAVAPADTP 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSARWDAQORPAVASRRRPTTAGAAPLTAVAPADTP 120
QY 121 PVPDVDSRGAILRRQVNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVDSRGAILRRQVNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYRPLVNPVNAVGGVAISISFWPQTTPPTTPTSDVMSITSTDVRLVQPGI 240
Db 181 NYAQVRVARATIRYRPLVNPVNAVGGVAISISFWPQTTPPTTPTSDVMSITSTDVRLVQPGI 240
QY 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVLNSYNTNTPTGTALGLL 300
Db 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVLNSYNTNTPTGTALGLL 300
QY 301 DFALBEPFNLTGNTNTRVSYSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
Db 301 DFALBEPFNLTGNTNTRVSYSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTLISAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTLISAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGTAIPHDIIDLGESRVVIQDYDQHEQDRPTSPAPSRPFSVLRANDVWLSLTAABY 480
Db 421 QDKGTAIPHDIIDLGESRVVIQDYDQHEQDRPTSPAPSRPFSVLRANDVWLSLTAABY 480
QY 481 DOSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTTQQYSKTFPVL 540
Db 481 DOSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTTQQYSKTFPVL 540
QY 541 LRGKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRGKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHASALALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGTREL 660
Db 601 LAPHASALALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGTREL 660
```

Search completed: August 16, 2004, 13:36:07

Job time : 21 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 13:35:07 ; Search time 51 Seconds
(without alignments)
4062.587 Million cell updates/sec

Title: US-09-851-410A-8
Perfect score: 3408
Sequence: 1 MRPRILLLLMLPMLPAP.....QSTVAELQRLKMKVGKTRRL 660

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	3408	100.0	660	9	US-09-769-066-13
2	3408	100.0	660	10	US-09-851-410-8
3	3408	100.0	660	14	US-10-165-868-19
4	3391	99.3	660	12	US-10-381-770-1
5	3383	99.3	660	15	US-10-257-044-1
6	3221	94.5	660	9	US-09-769-066-14
7	3194.5	93.7	659	14	US-10-165-868-20
8	3179	93.3	674	16	US-10-239-090A-57
9	3176	93.2	660	16	US-10-239-090A-51
10	3174	93.1	660	16	US-10-239-090A-53
11	3166	92.9	674	16	US-10-239-090A-55
12	3160	92.7	660	10	US-09-468-147-92
13	3160	92.7	660	12	US-10-319-745-92
14	3151	92.5	660	10	US-09-468-147-167
15	3151	92.5	660	12	US-10-319-745-167

16	2813	82.5	549	9	US-09-769-066-15	Sequence 15, Appl
17	2769	81.2	540	9	US-09-769-066-25	Sequence 25, Appl
18	2694	79.0	525	9	US-09-769-066-27	Sequence 27, Appl
19	2684	78.8	549	9	US-09-769-066-16	Sequence 16, Appl
20	2644	77.6	540	9	US-09-769-066-26	Sequence 26, Appl
21	2569	75.4	525	9	US-09-769-066-28	Sequence 28, Appl
22	2232	65.5	436	14	US-10-165-868-17	Sequence 17, Appl
23	2098.5	61.6	435	14	US-10-165-868-18	Sequence 18, Appl
24	1677	49.2	327	9	US-09-769-066-17	Sequence 17, Appl
25	1677	49.2	327	14	US-10-165-868-15	Sequence 15, Appl
26	1615	47.4	459	10	US-09-468-147-206	Sequence 206, App
27	1615	47.4	459	10	US-09-468-147-207	Sequence 207, App
28	1615	47.4	459	12	US-10-319-745-206	Sequence 206, App
29	1615	47.4	459	12	US-10-319-745-207	Sequence 207, App
30	1601	47.0	338	10	US-09-468-147-199	Sequence 199, App
31	1601	47.0	338	10	US-09-468-147-200	Sequence 200, App
32	1601	47.0	338	12	US-10-319-745-199	Sequence 199, App
33	1601	47.0	338	12	US-10-319-745-200	Sequence 200, App
34	1594	46.8	327	9	US-09-769-066-18	Sequence 18, Appl
35	1594	46.8	327	14	US-10-165-868-16	Sequence 16, Appl
36	1592	46.7	327	10	US-09-468-147-176	Sequence 176, App
37	1592	46.7	327	12	US-10-319-745-176	Sequence 176, App
38	1581	46.4	327	10	US-09-468-147-175	Sequence 175, App
39	1581	46.4	327	12	US-10-319-745-175	Sequence 175, App
40	1352	39.7	276	10	US-09-468-147-189	Sequence 189, App
41	1352	39.7	276	12	US-10-319-745-189	Sequence 189, App
42	1334.5	39.2	606	14	US-10-029-840-6	Sequence 6, Appli
43	1273	37.4	389	12	US-10-381-770-4	Sequence 4, Appli
44	1084	31.8	232	12	US-10-381-770-12	Sequence 12, Appl
45	1076	31.6	214	12	US-10-381-770-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-769-066-13
; Sequence 13, Application US/09769066
; Patent No. US20020107360A1

GENERAL INFORMATION:

APPLICANT: Fuerst, Thomas R.
McAtee, C. Patrick
Yarbough, Patrice O.
Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/769,066
FILING DATE: 24-Jan-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/542,634
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 13:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; STRANDEDNESS: Hepatitis E Virus (Burma strain)
; ORF-2
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-769-066-13

Query Match 100.0%; Score 3408; DB 9; Length 660;
Best Local Similarity 100.0%; Pred. No. 7.9e-276;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLFPLPAPPQPSGRRRGRSGSGGGFWGDRVDSQFFAIPYIHPTN 60
DB 1 MRPRPILLLLMLFPLPAPPQPSGRRRGRSGSGGGFWGDRVDSQFFAIPYIHPTN 60

QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPAHDT 120
DB 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPAHDT 120

QY 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGCTNLVLYAAPLSPLLPLODGTNTHIMATEAS 180
DB 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGCTNLVLYAAPLSPLLPLODGTNTHIMATEAS 180

QY 181 NYAQYRVARATIRYRPLVPNAVGGYALISFWPQTTTTPTSDVMNSITSDVRLVQPGI 240
DB 181 NYAQYRVARATIRYRPLVPNAVGGYALISFWPQTTTTPTSDVMNSITSDVRLVQPGI 240

QY 241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSLVLMCIHGSIVNSYTNPTYTGAIGLL 300
DB 241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSLVLMCIHGSIVNSYTNPTYTGAIGLL 300

QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
DB 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360

QY 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGGLFYRSPVWSANGEPPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGGLFYRSPVWSANGEPPTVKLYTSVENAQ 420

QY 421 QDRGIAIPHDIIDGESRVVLDYDQNHQEDRPTSPAPSPRPFVLRANDVLMLSLTAAY 480
DB 421 QDRGIAIPHDIIDGESRVVLDYDQNHQEDRPTSPAPSPRPFVLRANDVLMLSLTAAY 480

QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDRPLSTIOQYKSTFFVLP 540
DB 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDRPLSTIOQYKSTFFVLP 540

QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRAVISTYTTSLGAGPVSISAVV 600
DB 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRAVISTYTTSLGAGPVSISAVV 600

QY 601 LAPHSALEEDLTDYPARAHTDDFCPCRPGLGCAQOSTVAELQRLKMKVGTREL 660
DB 601 LAPHSALEEDLTDYPARAHTDDFCPCRPGLGCAQOSTVAELQRLKMKVGTREL 660
```

RESULT 2

```

US-09-851-410-8
; Sequence 8, Application US/09851410
; Publication No. US20030124510A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; Yarbough, Patrice O
; Bradley, Daniel W
; Krawczynski, Krzysztof Z
; Tam, Albert
; Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
```

```

; No. US20030124510A1-A/No. US20030124510A1-B Hepatitis Viru
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,410
; FILING DATE: 07-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/128,275
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-851-410-8
```

```

Query Match 100.0%; Score 3408; DB 10; Length 660;
Best Local Similarity 100.0%; Pred. No. 7.9e-276;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLFPLPAPPQPSGRRRGRSGSGGGFWGDRVDSQFFAIPYIHPTN 60
DB 1 MRPRPILLLLMLFPLPAPPQPSGRRRGRSGSGGGFWGDRVDSQFFAIPYIHPTN 60

QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPAHDT 120
DB 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPAHDT 120

QY 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGCTNLVLYAAPLSPLLPLODGTNTHIMATEAS 180
DB 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGCTNLVLYAAPLSPLLPLODGTNTHIMATEAS 180

QY 181 NYAQYRVARATIRYRPLVPNAVGGYALISFWPQTTTTPTSDVMNSITSDVRLVQPGI 240
DB 181 NYAQYRVARATIRYRPLVPNAVGGYALISFWPQTTTTPTSDVMNSITSDVRLVQPGI 240

QY 241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSLVLMCIHGSIVNSYTNPTYTGAIGLL 300
DB 241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSLVLMCIHGSIVNSYTNPTYTGAIGLL 300

QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
```

Db 301 DFALEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLPLTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLPLTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDLGSRVVIQYDQHQEODRTPSPAPSRPFSVLRLANDVLMSLTAAEY 480
Db 421 QDKGIAIPHDIIDLGSRVVIQYDQHQEODRTPSPAPSRPFSVLRLANDVLMSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQQYSKTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQQYSKTFFVLP 540
QY 541 LRKGLSFWAEGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWAEGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHASALALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660
Db 601 LAPHASALALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660

RESULT 3

US-10-165-868-19
; Sequence 19, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; Bradley, Daniel W.
; Twu, Jr-Shin
; Purdy, Michael A.
; Tam, Albert W.
; Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/165,868
; FILING DATE: 06-June-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-165-868-19

Query Match 100.0%; Score 3408; DB 14; Length 660;
Best Local Similarity 100.0%; Pred. No. 7.9e-276; Indels 0; Gaps 0;
Matches 660; Conservative 0; Mismatches 0;

QY 1 MRPRPILLLLMFLPMLPAPPQPSGRRRRSGSGGGFWGDRVDSQFFFAIPYIHPN 60
Db 1 MRPRPILLLLMFLPMLPAPPQPSGRRRRSGSGGGFWGDRVDSQFFFAIPYIHPN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAPLTAVAPAHDP 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAPLTAVAPAHDP 120
QY 121 PVPDVSRAILLRQYNLSTSLTSSVATGTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSRAILLRQYNLSTSLTSSVATGTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVENAVGGYASISFPQPTTTTPTSDVMSITSDVRILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVENAVGGYASISFPQPTTTTPTSDVMSITSDVRILVQPGI 240
QY 241 ASELVIPERLHYRQWRSVETSGVAEBEATSGLVMLCIHGSLSVNSYNTPTTGALGL 300
Db 241 ASELVIPERLHYRQWRSVETSGVAEBEATSGLVMLCIHGSLSVNSYNTPTTGALGL 300
QY 301 DFALEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTAATRFMKDLYFTSTNG 360
Db 301 DFALEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLPLTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLPLTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDLGSRVVIQYDQHQEODRTPSPAPSRPFSVLRLANDVLMSLTAAEY 480
Db 421 QDKGIAIPHDIIDLGSRVVIQYDQHQEODRTPSPAPSRPFSVLRLANDVLMSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQQYSKTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQQYSKTFFVLP 540
QY 541 LRKGLSFWAEGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWAEGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHASALALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660
Db 601 LAPHASALALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660

RESULT 4
US-10-381-770-1
; Sequence 1, Application US/10381770
; Publication No. US20040052813A1
; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPOS
; TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
; FILE REFERENCE: IEC010037PCT
; CURRENT APPLICATION NUMBER: US/10/381,770

; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 660
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-1

Query Match 99.5%; Score 3391; DB 12; Length 660;
Best Local Similarity 99.4%; Pred. No. 2.1e-274;
Matches 656; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY	1	MRPRPILLLMLPMLPAPPQPSGRRGRSGGGGFWGDRVDSQPPAIPYIHPTN	60
DB	1	MRPRPILLLMLPMLPAPPQPSGRRGRSGGGGFWGDRVDSQPPAIPYIHPTN	60
QY	61	PFAPDVTAAAGAGPRVQPARPLGSARWDOAQRPAVASRRRPTTAGAAPTAVAPADHTP	120
DB	61	PFAPDVTAAAGAGPRVQPARPLGSARWDOAQRPAVASRRRPTTAGAAPTAVAPADHTP	120
QY	121	PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS	180
DB	121	PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS	180
QY	181	NYAQYRVARATIRYRPLVNAVGVYAISISFWPQTTTTPTSDVMNSITSTDVRLVQPGI	240
DB	181	NYAQYRVARATIRYRPLVNAVGVYAISISFWPQTTTTPTSDVMNSITSTDVRLVQPGI	240
QY	241	ASELVIPSERLHYRNQWRSVETSGVAEEBATSGLVMLCIHGSVNSYTNTPYTGAJGLL	300
DB	241	ASELVIPSERLHYRNQWRSVETSGVAEEBATSGLVMLCIHGSVNSYTNTPYTGAJGLL	300
QY	301	DFALELFRNLTPGNTNTRVSRYSSTARHLRGGADGTAEITTTAAATFRMKDLYFTSTNG	360
DB	301	DFALELFRNLTPGNTNTRVSRYSSTARHLRGGADGTAEITTTAAATFRMKDLYFTSTNG	360
QY	361	VGEIGRGIALTLFNADTLGGLPTELISAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ	420
DB	361	VGEIGRGIALTLFNADTLGGLPTELISAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ	420
QY	421	QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSPFVSRLANDVWLSTAAEY	480
QY	481	DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDRPLSTIQOYSKTFFVLP	540
DB	481	DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDRPLSTIQOYSKTFFVLP	540
QY	541	LRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTVTTSLGAGPVSISAVV	600
DB	541	LRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTVTTSLGAGPVSISAVV	600
QY	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPLGQCAFGQSTVAELQRLKMKVGKTRREL	660
DB	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPLGQCAFGQSTVAELQRLKMKVGKTRREL	660

RESULT 5

US-10-257-044-1
; Sequence 1, Application US/10257044
; Publication No. US20030220475A1
; GENERAL INFORMATION:

; APPLICANT: The Government of the United States of America, as
; APPLICANT: Represented by the Secretary, Department of Health and
; APPLICANT: Human Services, Centers for Disease Control and Prevention
; APPLICANT: HOWARD A. FIELDS
; APPLICANT: YURI E. KHUDYAKOV
; APPLICANT: JIHONG MENG
; TITLE OF INVENTION: NEUTRALIZING IMMUNOGENIC HEV POLYPEPTIDES (HEV)
; FILE REFERENCE: 14114.034002
; CURRENT APPLICATION NUMBER: US/10/257,044
; CURRENT FILING DATE: 2002-10-04

; PRIOR APPLICATION NUMBER: 60/195,380
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Artificial Sequence
FEATURE:

; OTHER INFORMATION: Description of Artificial construct

; OTHER INFORMATION: Synthetic construct
US-10-257-044-1

Query Match 99.3%; Score 3383; DB 15; Length 660;
Best Local Similarity 99.4%; Pred. No. 9.7e-274;
Matches 656; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY	1	MRPRPILLLMLPMLPAPPQPSGRRGRSGGGGFWGDRVDSQPPAIPYIHPTN	60
DB	1	MRPRPILLLMLPMLPAPPQPSGRRGRSGGGGFWGDRVDSQPPAIPYIHPTN	60
QY	61	PFAPDVTAAAGAGPRVQPARPLGSARWDOAQRPAVASRRRPTTAGAAPTAVAPADHTP	120
DB	61	PFAPDVTAAAGAGPRVQPARPLGSARWDOAQRPAVASRRRPTTAGAAPTAVAPADHTP	120
QY	121	PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS	180
DB	121	PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS	180
QY	181	NYAQYRVARATIRYRPLVNAVGVYAISISFWPQTTTTPTSDVMNSITSTDVRLVQPGI	240
DB	181	NYAQYRVARATIRYRPLVNAVGVYAISISFWPQTTTTPTSDVMNSITSTDVRLVQPGI	240
QY	241	ASELVIPSERLHYRNQWRSVETSGVAEEBATSGLVMLCIHGSVNSYTNTPYTGAJGLL	300
DB	241	ASELVIPSERLHYRNQWRSVETSGVAEEBATSGLVMLCIHGSVNSYTNTPYTGAJGLL	300
QY	301	DFALELFRNLTPGNTNTRVSRYSSTARHLRGGADGTAEITTTAAATFRMKDLYFTSTNG	360
DB	301	DFALELFRNLTPGNTNTRVSRYSSTARHLRGGADGTAEITTTAAATFRMKDLYFTSTNG	360
QY	361	VGEIGRGIALTLFNADTLGGLPTELISAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ	420
DB	361	VGEIGRGIALTLFNADTLGGLPTELISAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ	420
QY	421	QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSPFVSRLANDVWLSTAAEY	480
DB	421	QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSPFVSRLANDVWLSTAAEY	480
QY	481	DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDRPLSTIQOYSKTFFVLP	540
DB	481	DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDRPLSTIQOYSKTFFVLP	540
QY	541	LRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTVTTSLGAGPVSISAVV	600
DB	541	LRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTVTTSLGAGPVSISAVV	600
QY	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPLGQCAFGQSTVAELQRLKMKVGKTRREL	660
DB	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPLGQCAFGQSTVAELQRLKMKVGKTRREL	660

RESULT 6

US-09-769-066-14
; Sequence 14, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:

; APPLICANT: Fuerst, Thomas R.
; McAttee, C. Patrick
; Yarbrough, Patrice O.
; Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31

```
/
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: Dehlinger & Associates
/ STREET: 350 Cambridge Ave., Suite 250
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94306
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ FILING DATE: 24-Jan-2001
/ CLASSIFICATION: <Unknown>
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/542,634
/ FILING DATE: <Unknown>
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 4600-0293.30
/ REFERENCE/DOCKET NUMBER: 4600-0293.30
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: Hepatitis E Virus (Mexico Strain)
/ ORF-2
/
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/
/ ORIGINAL SOURCE:
/ SEQUENCE DESCRIPTION: SEQ ID NO: 14:
/
/ US-09-769-066-14
/
/ Query Match 94.5%; Score 3221; DB 9; Length 660;
/ Best Local Similarity 93.5%; Pred. No. 3.4e-260;
/ Matches 617; Conservative 21; Mismatches 22; Indels 0; Gaps 0;
/
/ QY 1 MRPRPILLLLMLPLPAPPQPSRRRRSGSGGGFWDVDSQPFAPYTHPTN 60
/ DB 1 MRPRPILLLFLPLMPAPPTQPSRRRRSGSGTGGFWGDRVDSQPFAPYTHPTN 60
/
/ QY 61 PFAPDVTAAGAPRVQRPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPADHTP 120
/ DB 61 PFAPDVAAAGSGPRLRQPARPLGSTRDQAPRPSAASRRRPTAGAAALTAVAPADHTS 120
/
/ QY 121 PVEDVDSGAILRROYNLSTSLTSSVATGNLVLYAAPLPLPLQDGTNTHIMATEAS 180
/ DB 121 PVEDVDSGAILRROYNLSTSLTSSVATGNLVLYAAPLPLPLQDGTNTHIMATEAS 180
/
/ QY 181 NYAQVRVARATIRYRLPLVNAVGGVAISISFWPQTITTTTSVDMNSITSDVRILVQPGI 240
/ DB 181 NYAQVRVARATIRYRLPLVNAVGGVAISISFWPQTITTTTSVDMNSITSDVRILVQPGI 240
/
/ QY 241 ASELVIPSRLHYRNQGWRSVETSGVAEEATSGLVMLCIHGSVLNSYNTPTYGALGLL 300
/ DB 241 ASELVIPSRLHYRNQGWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTYGALGLL 300
/
/ QY 301 DPALELFNRLPGNTNTRVSYSSSTARHLRGGADGTAELTTAATRFMKDLYFTSTNG 360
/ DB 301 DPALELFNRLPGNTNTRVSYSSSTARHLRGGADGTAELTTAATRFMKDLYFTSTNG 360
/
/ QY 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
/ DB 361 VGEVGRGIALTLNLADTLGLLPTLISAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
/
/ QY 421 QDKGVAIPHDIIDLGSRVVIQDYDNQHQEDRTPSPAPSRPFSVLRANDVLMLSLTAABY 480
/ DB 421 QDKGVAIPHDIIDLGSRVVIQDYDNQHQEDRTPSPAPSRPFSVLRANDVLMLSLTAABY 480
/
/ QY 481 DOSTYGSSTGPVYSDSVTLNVNATCAQAVARSLDWTKVTLQGRPLSTTQOYSKTFVLP 540
/ DB 481 DOSTYGSSTGPVYSDSVTLNVNATCAQAVARSLDWTKVTLQGRPLSTTQOYSKTFVLP 540
/
/ QY 541 LRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAGHRVAISTYTTSLGAGPVYSISAVAV 600
/ DB 541 LRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAGHRVAISTYTTSLGAGPVYSISAVAV 600
/
/ QY 601 LAPHSALALLEDTLYPARAHTFDFCPCRPGLGQCAFQSTVABLQRLKMKVGTREL 660
/ DB 601 LAPRSALALLEDTFDPGHAHTFDFCPCRCALGLQCGAFQSTVABLQRLKMKVGTREL 660
/
/ RESULT 7
/ US-10-165-868-20
/ Sequence 20, Application US/10165868
/ Publication No. US20030143241A1
/ GENERAL INFORMATION:
/ APPLICANT: Reyes, Gregory R.
/ Twu, Jr-Shin
/ Purdy, Michael A.
/ Tam, Albert W.
/ Kwaczynski, Krzysztof Z.
/
/ TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dehlinger & Associates
/ STREET: 350 Cambridge Avenue, Suite 250
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94306
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ FILING DATE: 06-Jun-2002
/ CLASSIFICATION: <Unknown>
/ APPLICATION NUMBER: US/10/165,868
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/870,985A
/ FILING DATE: 20-APRIL-1992
/ APPLICATION NUMBER: US 822,335
/ FILING DATE: 17-JAN-1992
/ APPLICATION NUMBER: US 505,888
/ FILING DATE: 05-APRIL-1990
/ APPLICATION NUMBER: US 420,921
/ FILING DATE: 13-OCTOBER-1989
/ APPLICATION NUMBER: US 367,486
/ FILING DATE: 16-JUNE-1989
/ APPLICATION NUMBER: US 336,672
/ FILING DATE: 11-APRIL-1989
/ APPLICATION NUMBER: US 208,997
/ FILING DATE: 17-JUNE-1988
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0093.30
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 659 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/
```


Best Local Similarity 92.11%; Pred. No. 2e-256;
Matches 607; Conservative 26; Mismatches 26; Indels 0; Gaps 0;

```
QY 1 MRPRILLLLMLFPLPAPPPQSGRRRRSGGGGFWGDRVDSQPFAIPYIHPTN 60
Db 1 MRPRALLLLFLVLLPMLPAPAGQPSGRRRRSGGTGGGFWGDRVDSQPFAIPYIHPTN 60
QY 61 PFAPDVTAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPADTP 120
Db 61 PFAADVSPQAGACTPRQPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPADTPA 120
QY 121 PVPDVSRGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLOGDGTNTHIMATEAS 180
Db 121 PVPDVSRGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLOGDGTNTHIMATEAS 180
QY 181 NYAQVRARATIRYRPLVNAVGGYAIISIFWPQTTTPTSVDMSIITSTDVRIIVQPGI 240
Db 181 NYAQVRARATIRYRPLVNAVGGYAIISIFWPQTTTPTSVDMSIITSTDVRIIVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSVETSGVAREEATSGLVMLCIHGSLVNSYNTPTTCALGLL 300
Db 241 ASELVIPSERLHYRNOGWSVETSGVAREEATSGLVMLCIHGSLVNSYNTPTTCALGLL 300
QY 301 DPALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELTTTAAATRFMKDLYFTSTNG 360
Db 301 DPALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELTTTAAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLSSAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
Db 361 VGEVGRGIALTLFNLADTLGLLPTLSSAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
QY 421 QDKGIAI PHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
Db 421 QDKGITI PHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
QY 481 DQSTYGSSTGPPVYSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQYYSKTFYVLP 540
Db 481 DQTYGSSTNPMVYSDSVTLNVNATGAQAVARSLDWSKVTLDRPLTTIQYYSKTFYVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVGV 600
QY 601 LAPHSALEDTLDYPARAHTFDDFCPCRTLGLQGCAFSQSTVAELQRLKMKVKTR 659
Db 601 LAPHSALEDTLDYPARAHTFDDFCPCRTLGLQGCAFSQSTVAELQRLKMKVKTR 659
```

RESULT 10
US-10-239-090A-53
; Sequence 53, Application US/10239090A
; Publication No. US20040101820A1
; GENERAL INFORMATION:
; APPLICANT: KABUSHIKI KAISHA TOSHIBA
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co
; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same,
; TITLE OF INVENTION: detecting hepatitis E virus using the same
; FILE REFERENCE: 0280741P
; CURRENT APPLICATION NUMBER: US/10/239,090A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 57
; SEQ ID NO 53
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus JMY-Haw (ORF2)
US-10-239-090A-53

Query Match 93.11%; Score 3174; DB 16; Length 660;
Best Local Similarity 92.11%; Pred. No. 2.9e-256;
Matches 607; Conservative 25; Mismatches 27; Indels 0; Gaps 0;

```
QY 1 MRPRILLLLMLFPLPAPPPQSGRRRRSGGGGFWGDRVDSQPFAIPYIHPTN 60
Db 1 MRPRALLLLFLVLLPMLPAPAGQPSGRRRRSGGTGGGFWGDRVDSQPFAIPYIHPTN 60
QY 61 PFAPDVTAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPADTP 120
Db 61 PFAADVSPQAGACTPRQPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPADTPA 120
QY 121 PVPDVSRGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLOGDGTNTHIMATEAS 180
Db 121 PVPDVSRGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLOGDGTNTHIMATEAS 180
QY 181 NYAQVRARATIRYRPLVNAVGGYAIISIFWPQTTTPTSVDMSIITSTDVRIIVQPGI 240
Db 181 NYAQVRARATIRYRPLVNAVGGYAIISIFWPQTTTPTSVDMSIITSTDVRIIVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSVETSGVAREEATSGLVMLCIHGSLVNSYNTPTTCALGLL 300
Db 241 ASELVIPSERLHYRNOGWSVETSGVAREEATSGLVMLCIHGSLVNSYNTPTTCALGLL 300
QY 301 DPALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELTTTAAATRFMKDLYFTSTNG 360
Db 301 DPALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELTTTAAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLSSAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
Db 361 VGEVGRGIALTLFNLADTLGLLPTLSSAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
QY 421 QDKGIAI PHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
Db 421 QDKGITI PHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
QY 481 DQSTYGSSTGPPVYSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQYYSKTFYVLP 540
Db 481 DQTYGSSTNPMVYSDSVTLNVNATGAQAVARSLDWSKVTLDRPLTTIQYYSKTFYVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVGV 600
QY 601 LAPHSALEDTLDYPARAHTFDDFCPCRTLGLQGCAFSQSTVAELQRLKMKVKTR 659
Db 601 LAPHSALEDTLDYPARAHTFDDFCPCRTLGLQGCAFSQSTVAELQRLKMKVKTR 659
```

RESULT 11
US-10-239-090A-55
; Sequence 55, Application US/10239090A
; Publication No. US20040101820A1
; GENERAL INFORMATION:
; APPLICANT: KABUSHIKI KAISHA TOSHIBA
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co
; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same, i
; TITLE OF INVENTION: detecting hepatitis E virus using the same
; FILE REFERENCE: 0280741P
; CURRENT APPLICATION NUMBER: US/10/239,090A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 57
; SEQ ID NO 55
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Hepatitis E Virus JKK-Sap (ORF2)
US-10-239-090A-55

Query Match 92.9%; Score 3166; DB 16; Length 674;
Best Local Similarity 92.0%; Pred. No. 1.4e-255;
Matches 606; Conservative 24; Mismatches 29; Indels 0; Gaps 0;

```
QY 1 MRPRILLLLMLFPLPAPPPQSGRRRRSGGGGFWGDRVDSQPFAIPYIHPTN 60
Db 15 MRSRAFLFLVLLPMLPAPAGQPSGRRRRSGAGGFWGDRVDSQPFAIPYIHPTN 74
```

QY 61 PFAPDVTAAAGAGRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPTAVAPADHTP 120
Db 75 PFASDTPAAAGAGRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPTAVAPADHTP 134
QY 121 PVPDVSRRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLPLQDGTNTHIMATEAS 180
Db 135 PVPDVSRRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLPLQDGTNTHIMATEAS 194
QY 181 NYAQYRVARATIRYRPLVNAVGGYAIISIFWFPQTTTTPTSDVMNSITSDVRILVQPGI 240
Db 195 NYAQYRVARATIRYRPLVNAVGGYAIISIFWFPQTTTTPTSDVMNSITSDVRILVQPGI 254
QY 241 ASELVIPSERLHYRNQWRSVETSGVAEEATSGLVMLCTHGSIVNSYTNTPYTGALL 300
Db 255 ASELVIPSERLHYRNQWRSVETSGVAEEATSGLVMLCTHGSIVNSYTNTPYTGALL 314
QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 315 DFALLEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 374
QY 361 VGEIGRGIALTLFNADTLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 375 VGEVGRGIALTLFNADTLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 434
QY 421 QDKGIALPHDIDGESRVVIQDYNQHEODRPTSPAPSPFPFVLRANDVWLSTAAEY 480
Db 435 QDKGIALPHDIDGESRVVIQDYNQHEODRPTSPAPSPFPFVLRANDVWLSTAAEY 494
QY 481 DQSTYSSGTPVYVSDSVTLNVNATGAQAVASLDWTKVLDGRPLSTIOQYKSTFFVLP 540
Db 495 DQSTYSSGTPVYVSDSVTLNVNATGAQAVASLDWTKVLDGRPLSTIOQYKSTFFVLP 554
QY 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAHRAVIASTYTTSLGAGPVSISAVV 600
Db 555 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAHRAVIASTYTTSLGAGPVSISAVV 614
QY 601 LAPHSALELLEDVDPARAHTFDDFCPECRPLGQCAFQSTVAELQRLKMKVGTRE 659
Db 615 LAPHSALELLEDVDPARAHTFDDFCPECRPLGQCAFQSTVAELQRLKMKVGTRE 673

RESULT 12

US-10-319-745-92
; Sequence 92, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/09/468,147A
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/173,141
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 92
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 481
; OTHER INFORMATION: Xaa = Unknown or Other at position 542
; OTHER INFORMATION: Xaa = Unknown or Other at position 595

US-09-468-147-92

Query Match 92.7%; Score 3160; DB 10; Length 660;
Best Local Similarity 91.7%; Pred. No. 4.3e-255;
Matches 604; Conservative 27; Mismatches 28; Indels 0; Gaps 0;
QY 1 MRPRPILLLMLPMLPAPPQPSGRRRGRSGSGGFWGDRVDSQPPAIPYIHTN 60
Db 1 MRPRPILLLMLPMLPAPPQPSGRRRGRSGSGGFWGDRVDSQPPAIPYIHTN 60
QY 61 PFAPDVTAAAGAGRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPTAVAPADHTP 120
Db 61 PFAPDVTAAAGAGRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPTAVAPADHTP 120
QY 121 PVPDVSRRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVNAVGGYAIISIFWFPQTTTTPTSDVMNSITSDVRILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGGYAIISIFWFPQTTTTPTSDVMNSITSDVRILVQPGI 240
QY 241 ASELVIPSERLHYRNQWRSVETSGVAEEATSGLVMLCTHGSIVNSYTNTPYTGALL 300
Db 241 ASELVIPSERLHYRNQWRSVETSGVAEEATSGLVMLCTHGSIVNSYTNTPYTGALL 300
QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNADTLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEVGRGIALTLFNADTLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIALPHDIDGESRVVIQDYNQHEODRPTSPAPSPFPFVLRANDVWLSTAAEY 480
Db 421 QDKGIALPHDIDGESRVVIQDYNQHEODRPTSPAPSPFPFVLRANDVWLSTAAEY 480
QY 481 DQSTYSSGTPVYVSDSVTLNVNATGAQAVASLDWTKVLDGRPLSTIOQYKSTFFVLP 540
Db 481 DQSTYSSGTPVYVSDSVTLNVNATGAQAVASLDWTKVLDGRPLSTIOQYKSTFFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAHRAVIASTYTTSLGAGPVSISAVV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAHRAVIASTYTTSLGAGPVSISAVV 600
QY 601 LAPHSALELLEDVDPARAHTFDDFCPECRPLGQCAFQSTVAELQRLKMKVGTRE 659
Db 601 LAPHSALELLEDVDPARAHTFDDFCPECRPLGQCAFQSTVAELQRLKMKVGTRE 659

RESULT 13

US-10-319-745-92
; Sequence 92, Application US/10319745
; Publication No. US20030211467A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/10/319,745
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: US/09/468,147A
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/173,141
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/061,199
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 92
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 481
; OTHER INFORMATION: Xaa = Unknown or Other at position 542
; OTHER INFORMATION: Xaa = Unknown or Other at position 595
US-10-319-745-92
```

```
Query Match          92.7%; Score 3160; DB 12; Length 660;
Best Local Similarity 91.7%; Pred. No. 4.3e-255;
Matches 604; Conservative 27; Mismatches 28; Indels 0; Gaps 0;
```

```
QY 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGFWGDRVDSQFFAIPYIHPN 60
Db 1 MRPRVLLLLFVLLPMLPAPPAGQPSGRRGRSGSGGFWGDRVDSQFFAIPYIHPN 60
QY 61 PFAPDVTAAGAGPRVRQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADTP 120
Db 61 PFAADVSPQAGTRPRQPPRLGSAWRDQSKPSVAPRRSTPAGAAPLTAVSPADTA 120
QY 121 PVPDVDSGAILRRQVNLSTPLTSSVATGTLNLYAAPLSPLPLQDGTNTHIMATEAS 180
Db 121 PVPDVDSGAILRRQVNLSTPLTSSVATGTLNLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVNAVGVYASISFWPQTTTTTSDVMSNITSDVRILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGVYASISFWPQTTTTTSDVMSNITSDVRILVQPGI 240
QY 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGTALGLL 300
Db 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGTALGLL 300
QY 301 DPALAEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
Db 301 DPALAEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLGLPTELISSAGQLFYSPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLGLPTELISSAGQLFYSPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGTAIPHDIIDGSRVVIQDYDQHEQDRTPSPAPSRPFSVLRANDVLSLTAABY 480
Db 421 QDKGTAIPHDIIDGSRVVIQDYDQHEQDRTPSPAPSRPFSVLRANDVLSLTAABY 480
QY 481 DQSTYGSSTGPPVYVSDVTLNVNATGAQAVARSLDWTKVTLDGRPLSTTIQQYSKTFFVLP 540
Db 481 DQSTYGSSTGPPVYVSDVTLNVNATGAQAVARSLDWTKVTLDGRPLSTTIQQYSKTFFVLP 540
QY 541 LRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDTLDYPARAHTFDDFCPCRLGLGQCAFQSTVAELQRLKMKVKTRE 659
Db 601 LAPHSALELLEDTLDYPARAHTFDDFCPCRLGLGQCAFQSTVAELQRLKMKVKTRE 659
```

RESULT 14

```
US-09-468-147-167
; Sequence 167, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauer, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushanwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
```

```
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/09/468,147A
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/173,141
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 167
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 84
; OTHER INFORMATION: Xaa = Unknown or Other at position 230
; OTHER INFORMATION: Xaa = Unknown or Other at position 447
US-09-468-147-167
```

```
Query Match          92.5%; Score 3151; DB 10; Length 660;
Best Local Similarity 91.5%; Pred. No. 2.4e-254;
Matches 603; Conservative 26; Mismatches 30; Indels 0; Gaps 0;
```

```
QY 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGFWGDRVDSQFFAIPYIHPN 60
Db 1 MRPRVLLLLFVLLPMLPAPPAGQPSGRRGRSGSGGFWGDRVDSQFFAIPYIHPN 60
QY 61 PFAPDVTAAGAGPRVRQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADTP 120
Db 61 PFAADVSPQAGTRPRQPPRLGSAWRDQSKPSVAPRRSTPAGAAPLTAVSPADTA 120
QY 121 PVPDVDSGAILRRQVNLSTPLTSSVATGTLNLYAAPLSPLPLQDGTNTHIMATEAS 180
Db 121 PVPDVDSGAILRRQVNLSTPLTSSVATGTLNLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVNAVGVYASISFWPQTTTTTSDVMSNITSDVRILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGVYASISFWPQTTTTTSDVMSNITSDVRILVQPGI 240
QY 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGTALGLL 300
Db 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGTALGLL 300
QY 301 DPALAEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
Db 301 DPALAEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLGLPTELISSAGQLFYSPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLGLPTELISSAGQLFYSPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGTAIPHDIIDGSRVVIQDYDQHEQDRTPSPAPSRPFSVLRANDVLSLTAABY 480
Db 421 QDKGTAIPHDIIDGSRVVIQDYDQHEQDRTPSPAPSRPFSVLRANDVLSLTAABY 480
QY 481 DQSTYGSSTGPPVYVSDVTLNVNATGAQAVARSLDWTKVTLDGRPLSTTIQQYSKTFFVLP 540
Db 481 DQSTYGSSTGPPVYVSDVTLNVNATGAQAVARSLDWTKVTLDGRPLSTTIQQYSKTFFVLP 540
QY 541 LRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDTLDYPARAHTFDDFCPCRLGLGQCAFQSTVAELQRLKMKVKTRE 659
Db 601 LAPHSALELLEDTLDYPARAHTFDDFCPCRLGLGQCAFQSTVAELQRLKMKVKTRE 659
```

RESULT 15

```
US-10-319-745-167
; Sequence 167, Application US/10319745
; Publication No. US20030211467A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/10/319,745
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: US/09/468,147A
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/173,141
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/061,199
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 167
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 84
; OTHER INFORMATION: Xaa = Unknown or Other at position 230
; OTHER INFORMATION: Xaa = Unknown or Other at position 447
US-10-319-745-167
```

Search completed: August 16, 2004, 13:41:00
Job time : 52 secs

Qy 601 LAPHSALELLELTLDYPARAHTDDFCPCRPGLGQCAFSQSTVAELQRLKMKVGKTR 659
Db 601 LAPHSALELLELTLDYPARAHTDDFCPCRPGLGQCAFSQSTVAELQRLKMKVGKTR 659

```
Query Match 92.5%; Score 3151; DB 12; Length 660;
Best Local Similarity 91.5%; Pred. No. 2.4e-254;
Matches 603; Conservative 26; Mismatches 30; Indels 0; Gaps 0;

Qy 1 MRPRPILLLLMLPAPPQPSGRRRGRSGGSGGFWGDRVDSQPFAPYIHPN 60
Db 1 MRPRVLLLLFVLLPMLPAPPQPSGRRRGRSGGSGGFWGDRVDSQPFAPYIHPN 60

Qy 61 PRAPDVTAAAGAPRVQAPRLGSAWEDQAPAVASRRPTTAGAAPTAVAPADHTP 120
Db 61 PFAADVVSQAGAGPRPQPRPLXSAWEDQAPAVASRRPTTAGAAPTAVAPADHTP 120

Qy 121 PVPDVSRGAILRRQYNLSPTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSRGAILRRQYNLSPTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180

Qy 181 NYAQYRVARATIRYRPLVPNAVGYAISISFWPQTTTTPTSDVMNSITSTDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRPLVPNAVGYAISISFWPQTTTTPTSDVMNSITSTDVRLVQPGI 240

Qy 241 ASELVIPSERLHYRNOGWRSEVTSVGAEEATSGLVMLCIHGSIVNSYTNTPYTGALGLL 300
Db 241 ASELVIPSERLHYRNOGWRSEVTSVGAEEATSGLVMLCIHGSIVNSYTNTPYTGALGLL 300

Qy 301 DFALALEFRNLTPGNTNTRYSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALALEFRNLTPGNTNTRYSRYSSTARHRLRGADGTAEITTTAATRFMKDLHFGING 360

Qy 361 VGEIGRGIALTLNLAADTLGGPTELISAGGOLFYSRPPVVSANGEPYVKLYTSVENAQ 420
Db 361 VGEVGRGIALTLNLAADTLGGPTELISAGGOLFYSRPPVVSANGEPYVKLYTSVENAQ 420

Qy 421 QDKGIATPHDIDIGESRVVICDYNQHEQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
Db 421 QDKGITPHDIDIGDSRWIQDYNQHEQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480

Qy 481 DQSTYGSSTGPEVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQQYKTFVLP 540
Db 481 DQTYGSSITNPMYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQQYKTFVLP 540

Qy 541 LRKGLSFEAGTTKAGYPYNYNTTASDQILLIENAAAGHRVAISTYTTSLGAGPVSISAVV 600
Db 541 LRKGLSFEAGTTKAGYPYNYNTTASDQILLIENAAAGHRVAISTYTTSLGAGPVSISAVV 600
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: August 16, 2004, 13:35:43 ; Search time 89.8179 Seconds
(without alignments)
12476.164 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129
Perfect score: 660
Sequence: 1 ATGCGCCTCGGCTATTTT.....GTAAACTCGGAGTTGTAG 1983

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2987630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xl
-Q=/Cgn2_1/USPTO_spool_p/US09851410/runat_16082004_125716_9617/app_query.fasta_1.4238
-DB=A Geneseq_29Jan04 -QWMT=fastan -SUFFIX=oli.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09851410 @CGN_1_1_141@runat_16082004_125716_9617 -NCPU=6 -ICPU=3
-NO_WMAP -LARGSEQURY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	660	100.0	660	2 AAR38785	Aar38785 HEV ORF2
2	660	100.0	660	2 AAR39306	Aar39306 Burma str
3	660	100.0	660	2 AAR96089	Aar96089 Hepatitis
4	660	100.0	660	2 AAW35826	Aaw35826 Hepatitis
5	660	100.0	660	2 AAW80197	Aaw80197 Protein e
6	660	100.0	660	2 AAW93389	Aaw93389 Human HEV
7	660	100.0	660	3 AAB24120	Aab24120 Hepatitis
8	660	100.0	660	4 AAB62523	Aab62523 HEV-Burma
9	660	100.0	660	5 AAO15699	Aao15699 Hepatitis
10	660	100.0	660	7 ADD24376	Add24376 Hepatitis

11	559	84.7	660	2 AAW76369	Aaw76369 Hepatitis
12	559	84.7	660	2 AAW71210	Aaw71210 Protein e
13	549	83.2	549	2 AAR96091	Aar96091 Hepatitis
14	541	82.0	660	4 AAE12980	Aae12980 Hepatitis
15	540	81.8	540	2 AAR96101	Aar96101 Hepatitis
16	525	79.5	525	2 AAR96103	Aar96103 Hepatitis
17	512	77.6	660	2 AAR70323	Aar70323 Hepatitis
18	448	67.9	549	2 AAW76367	Aaw76367 Hepatitis
19	446	67.6	660	5 ABJ04803	Abj04803 Hepatitis
20	408	61.8	660	2 AAR91814	Aar91814 Hepatitis
21	408	61.8	660	2 AAW93386	Aaw93386 Human HEV
22	408	61.8	660	5 ABB81670	Abb81670 Hepatitis
23	408	61.8	660	6 ADA50060	Ada50060 Hepatitis
24	408	61.8	660	7 ADB97807	Adb97807 HEV ORF2
25	396	60.0	660	2 AAW93388	Aaw93388 Human HEV
26	327	49.5	327	2 AAR96093	Aar96093 Hepatitis
27	327	49.5	327	2 AAW35820	Aaw35820 Hepatitis
28	321	48.6	660	2 AAR51265	Aar51265 HEV strai
29	321	48.6	660	2 AAW81520	Aaw81520 Hepatitis
30	310	47.0	660	5 ABJ04816	Abj04816 Hepatitis
31	304	46.1	660	5 ABJ04820	Abj04820 Hepatitis
32	298	45.2	660	5 ABJ04811	Abj04811 Hepatitis
33	284	43.0	660	2 AAW93395	Aaw93395 Human HEV
34	264	40.0	660	2 AAW93392	Aaw93392 Human HEV
35	259	39.2	660	2 AAW93390	Aaw93390 Human HEV
36	255	38.6	660	5 ABJ04807	Abj04807 Hepatitis
37	251	38.0	660	5 ABJ04818	Abj04818 Hepatitis
38	251	38.0	660	5 ABJ04809	Abj04809 Hepatitis
39	250	37.9	660	5 ABJ04819	Abj04819 Hepatitis
40	228	34.5	660	2 AAR14619	Aar14619 Protein e
41	217	32.9	660	5 ABJ04822	Abj04822 Hepatitis
42	214	32.4	660	5 ABJ04821	Abj04821 Hepatitis
43	210	31.8	213	4 AAU00496	Aau00496 Hepatitis
44	210	31.8	214	5 ABB81671	Abb81671 Hepatitis
45	210	31.8	232	5 ABB81679	Abb81679 Hepatitis

ALIGNMENTS

RESULT 1
AAR38785
ID AAR38785 standard; protein; 660 AA.
XX
AC AAR38785;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 11-JAN-1994 (first entry)
XX
DE HEV ORF2 protein.
XX
KW Enterically transmitted non-A non-B hepatitis; ET-NANB; vaccine.
XX
OS Hepatitis E virus; Burma strain.
XX
FH Key Location/Qualifiers
FT Peptide 225..660
FT Peptide 334..660
FT Peptide /label= SG3
FT Peptide 613..660
FT Peptide /label= 406.3-2
XX
PN WO9314116-A1.
XX
PD 22-JUL-1993.
XX
PF 15-JAN-1993; 93WO-US000459.
XX
PR 17-JAN-1992; 92US-00822335.
PR 01-MAY-1992; 92US-00876941.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.

```
PA (USSH ) US SEC DEPT HEALTH.
XX
PI Reyes GR, Bradley DW, Tam AW, Carl M;
XX
XX WPI; 1993-243144/30.
DR
DR N-PSDB; AAQ47129.
XX
XX New immunogenic hepatitis E virus (HEV) peptide(s) - are from the ORF1,
PT ORF2 and ORF3 regions of HEV, useful as a vaccine against HEV infection.
PS Disclosure; Fig 8; 48pp; English.
XX
XX Immunogenic hepatitis E virus (HEV) peptides are selected from the ORF1,
CC ORF2 and ORF3 regions of HEV. The peptides can be used in vaccines to
CC prevent infection by HEV. The antibodies can neutralise and block HEV
CC infection and can be used to prevent or treat HEV infection. The peptides
CC and antibodies can also be used as diagnostic reagents. (Updated on 25-
CC MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
XX
XX Sequence 660 AA;
SQ
Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAR38785 (1-660)
QY 1 ATCGCGCTCGGCTATTTGTCGCTCTCATGTTTGGCTATGTCGCGCGCCA 60
DB 1 MetArgProArgProLeuLeuLeuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCGCGGTGACGGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC 120
DB 21 ProProGlyGlnProSerGlyArgArgArgGlyArgSerGlyGlyGly 40
QY 121 TTTGGGGTACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCCCATGTCACGGTGGCGCGGGCTGGACCTCGTTCGCCAACCGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyValArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCCGCTTGGCGGTGACGAGCCCGCGCGCGTTCGCTCACGTCGT 300
DB 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACCTTACCAAGTGGGGCGGCGCTAACCGCGGTGCTCGGCGCCATGACACCCCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCTGATGTCGACTCCCGCGGCGCATCTTGGCGCGGAGTATACCTATCAACA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnIleAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCGGCGCACCGGCACTAACCTGGTCTCTTATGCGGCCCTCT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTTACCTTTAGGACGGACCAATACCCATATATGCGCAGGAGCTTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTTATGCCAGTACCGGGTTCGCGTGCACAAATCGTTTACGCCCGGTGGTCCCAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTGCGCGGTACGCCATCTCCATCTCATTTCTGGCCACAGACCACCCCGGACG 660
201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
661 TCGGTTGATGATGAATTCATTAACCTCGACGGATGTTGGTATTTTAGTCCAGCCGCGCATA 720
221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
721 GCTCTGAGCTTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGSCTCGCGCTCC 780
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
781 GTCCAGACCTCTCGGGTGGCTGAGGAGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
841 CATGGCTCACTCGTAATTCCTATATACTACACCTATACCGTCCGCTCGGGCTGTTG 900
281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeu 300
901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACTTACCCCGGTAAACCAATACGCGGGTC 960
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
961 TCCCGTTATTCAGCAGCTGCTCGCCACCGCTTCGTCGCGGTGCGGAGCGGAGTCCGAG 1020
321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
1021 CTCACCAACACGCTGCTACCGCTTATCAAGAGCTCTATTTTACTAGTACTAATGGT 1080
341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTGTGTGACACTCTGCTT 1140
361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
1141 GCGCGCTCCGACAGAAATGATTCGTGGCTGGTGGCGAGCTGTCTACTCCGTCCTCC 1200
381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
1201 GTTGTCTCAGCAATCGCGAGCGCTGTAACTGTATATCATCTGTAGAGATGCTCAG 1260
401 ValValSerAlaAsnGlyGluProThrValIleLeuTyrThrSerValGluAsnAlaGln 420
1261 CAGGATAAGGATTTGCAATCCCGCATGACATTGACCTCGGAGAATCTCGTGGTTATT 1320
421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
1321 CAGGATTTATGATAACCAACATGAACAGATCGCGCGAGCGCTTCTCCAGCCCATCGCGC 1380
441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
1381 CCTTCTCTGCTCTCGAGCTAATGATGTGCTTGGCTCTCTCTCAGCGTGGCGAGTAT 1440
461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGlyTyr 480
1441 GACAGTCCACTTATGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTG 1500
481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
1501 GTTAATGTTGCGACCGCGCGGAGCGGCTGCGCGGTGCGCTCGATTGGACCAAGTCCACA 1560
501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
1561 CTTGAGCGTCCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTCTGCCG 1620
521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
1621 CTCGCGGGTAAGCTCTCTTTCTCGGAGGCGAGGACCAACTAAAGCCGGGTACCTTATAAT 1680
541 LeuArgGlyLysLeuSerPheThrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
1681 TATAACCACTGCTAGCGACCACTGCTTGTGAGATCGCGCGGACCGCGGTGCT 1740
```

Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACCACTAGCTGGTGTCTGCTCCGCTCCATTTCTGCGGTTGCCGTT 1800
Db 581 IleSerThrTyThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTGCGCTAGCATTTGCTTGGAGTACCTTGGACTTACCTCCCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyProAlaArgAla 620
QY 1861 CATACTTTTGATGATTTCTGCCAGAGTCCCGCCCTTGGCTTTCAGGGCTGCGCTTC 1920
Db 621 HisThrPheAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGCTACTGCTCGCTGAGCTTCAGCGCTTACAGTGAAGTGGTAAACTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuMetLysValGlyLysThrArgGluLeu 660

RESULT 2
AAR39306
ID AAR39306 standard; protein; 660 AA.
XX
AC AAR39306;
XX
DT 25-MAR-2003 (revised)
DT 14-FEB-1994 (first entry)
XX
DE Burma strain HEV ORF2 putative virus capsid protein.
XX
KW Hepatitis E virus; vaccine; neutralising antibodies; infection; block;
KW open reading frame; antibodies.
XX
OS Hepatitis E virus.
XX
PN WO9314208-A2.
XX
PD 22-JUL-1993.
XX
PF 19-JAN-1993; 93WO-US000475.
XX
PR 17-JAN-1992; 92US-00822335.
PR 20-APR-1992; 92US-00870985.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Reyes GR, Bradley DW, Twu J, Purdy MA, Tam AW, Krawczynski KZ;
XX WPI; 1993-243223/30.
XX N-PSDB; AAQ46813.
XX
PT Antigen and antibody vaccines against hepatitis E virus infection -
PT contain peptide(s) derived from capsid protein C-terminal or antibodies
PT against protein.
XX
PS Disclosure; Fig 7; 43pp; English.
XX
CC The sequence is that of the putative virus capsid protein encoded by
CC Burma strain hepatitis E virus (HEV) open reading frame ORF2. This
CC protein or peptide fragments of it may be used in a vaccine composition
CC for immunising an individual against HEV. Antibodies raised against these
CC peptides can also be used in such vaccines. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 660 AA;

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-851-410a-6_COPY_5147_7129 (1-1983) x AAR39306 (1-660)
QY 1 ARGCCTCGGCTCATTTTGTGTGTCTCATGTCTTTTGCCTATGCTGCCGCGCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCCCGGCTCAGCGCTCTGCGCGCTCGTGGCGGCGCAGCGGCTTCCCGCGTGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly 40
QY 121 TCTCGGCTGACCGGCTTGAATCTCAGCGCTTCCCAATCCCTATATTCATCAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyIleHisProThrAsn 60
QY 181 CCCTTGGCCCCGATGTCAACGCTGCGCGCGGCTGAGCTCTGTGTCGCAACCCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCCGCTTGGCTGACGAGGCGCGCGCGCGCTTGCCTCAGCTGT 300
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACCTACACAGCTGGCGCGCGCTAACCGCGCTCGCTCCGCGCGCATGACACCCCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CAGTGCCTGATGCTGACTCCCGCGCGCTTCTTGGCGCGGCTGAGTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCTGGCGCACCGGCTAACTGCTTCTTTATGCGCGCTCTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTTACCCCTTCCAGGACGCGCAATACCATATATGCGCGCAAGCTTCT 540
Db 161 SerProLeuLeuProLeuGluAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGCTTCCCGTCCCAATCCGTTACCGCGCGCTGCTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGGGGTTAGCCCTCTCCATCTCATCTGCGCACAGACACACACCCCGAG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATTAACCTCGACGATGTTCTGTTATTTAGTCCAGCGCGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTGTAGCTGTGATCCCAAGTGAAGCTTACACTATCGTAAACCAAGGTGGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTTGGGGTGGCTGAGAGGAGGCTACTCTGCTTGTATTGCTTTGCTAT 840
Db 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCAGCTCGTAAATTCCTATATAATACACCTATACCGGTGCGCTCGGGCTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTCCTTGTAGCTTGTGCAACCTTACCCCGCTTACCAACATACCGGGTGC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGTTATTCAGCACTGCTCGCCACCGCTTGTGCGGTGGGAGGAGTGGCCGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCGCTGCTTACCGCTTTTATGAAGACCTCTATTATTTACTAGTACTAATCGT 1080

QY 361 CCAGTGCCTGATGTCGACTCCCGCGCGCCATCTTTGCGCGGAGTATAACCTATCAACA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuAArgGlnTyAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCTGCGCCACCGGCACTAACCTGGTTCCTTATGCGCCCTCTT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyAlaAlaProLeu 160
QY 481 AGTCGGCTTTACCTCCCTTCAGACGCGCACCAATACCATATATATATGCGCACGAACTTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGCTTCCCGCTGCCAATCCGTTACCGCCGCTGGTCCCAAT 600
DB 181 AsnTyAlaGlnTyArgValAlaArgAlaThrIleAArgTyArgProLeuValProAsn 200
QY 601 GCTGTGCGGGTTACGCCATCTCCATCTCATCTCTGCGCCACAGACACCAACCCGAGC 660
DB 201 AlaValGlyGlyTyAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATACCTCGAGGATGTCGTATTTAGTCAGCGCCGCATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTGTGATCCCAAGTCAGCGCTACACTATCGTAAACCAAGCTGGCGCTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyArgAsnGlnGlyTrpArgSer 260
QY 781 CTCGAGACCTCTGGGCTGAGGAGGAGGCTACCTCTGCTCTTGTATGCTTTGCATA 840
DB 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCACTCGTAAATTCCTATCTACTAATACACCTATACCGTCCCTCGGCTGTG 900
DB 281 HisGlySerLeuValAsnSerTyThrAsnThrProTyThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGACCTTGAGTTTCGCAACCTTACCCCGGTAACACCAATACGGGGTC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCTTTATTCAGCACTGCTGCCACCGCTTCGTCGCGGTGCGGAGCGGACTGCCGAG 1020
DB 321 SerArgTySerSerThrAlaArgHisArgLeuAArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACGCGCTACCCGCTTATGAGGACCTCTATTTACTAGTACTAATGCT 1080
DB 341 LeuThrThrThrAlaAlaThrArgPheMetIysAspLeuTyPheThrSerThrAsnGly 360
QY 1081 CTCGCTGAGATCGCGCGGATAGCCCTTCAACCTGTTCAACCTGCTGACACTCTGCTT 1140
DB 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCGACAGAAATGATTTGCTGCTGAGTGTGTCGCTGCTTCTTCTCCGTCGC 1200
DB 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTySerArgPro 400
QY 1201 GTTGTCTCAGCAATGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 401 ValValSerAlaAsnGlyGlnProThrValIysLeuTyThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAGGGTATTCGAATCCCGATGACATTCGAGTTCGAGTTCGCTGCTGCTGCTT 1320
DB 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGlnSerArgValIle 440
QY 1321 CAGGATATGATACCAACATCAAGATCGCGGACGCTTCTCCAGCCCCCTCGCGC 1380
DB 441 GlnAspTyAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTTCTCTCTCTCGAGCTAATGATGCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
DB 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTy 480
QY 1441 GACCAGTCCACTTATGGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500

DB 481 AspGlnSerThrTyThrGlySerSerThrGlyProValTyValSerAspSerValThrLeu 500
QY 1501 GTTAAATGTTGACACCGCGCGGAGCGCTTCCCGCTGCTCGATTCGACCAAGGTCA 1560
DB 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaAArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTTGACGGTCCGCCCTCTCCACCATCCAGAGTACTCGAAGACCTCTTTGCTCCGCG 1620
DB 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTySerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAAAGCTCTCTTTCTGGGAGGACGACCACTAAAGCGCGGTACCTTATAAT 1680
DB 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyProTyAsn 560
QY 1681 TATAACACACTGCTAGGACCACTGCTTCGAGATGCGCGCGGACCGGTCGCT 1740
DB 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACCACTAGCTGGTGTGCTGCTCCGCTCTCCATTTCTGCGTTGCCGT 1800
DB 581 IleSerThrTyThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGAGTACCTGCCCGCGC 1860
DB 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyProAlaArgAla 620
QY 1861 CATACTTTGATGATTTCTGCCAGAGTGCGCCCTTGGCTTCAGGCTTCGCGCTTC 1920
DB 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAGTGAAGTGGTAAACTCGGAGTTG 1980
DB 641 GlnSerThrValAlaGluLeuGlnArgLeuIysMetLysValGlyIysThrArgGluLeu 660

RESULT 4
AAW35826
ID AAW35826 standard; protein; 660 AA.
XX AAW35826;
XX
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 26-FEB-1998 (first entry)
XX
DE Hepatitis E virus Burma strain protein from ORF2.
XX
KW Hepatitis E virus; Burma; Mexico; immunoassay; peptide antigen; antibody; diagnosis; HEV.
XX
OS Hepatitis E virus; - Burma strain.
XX
PN US5686239-A.
XX
PD 11-NOV-1997.
XX
PF 09-MAY-1994; 94US-00240049.
XX
PR 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-00336672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 05-APR-1991; 91US-00681078.
PR 17-JAN-1992; 92US-00822335.
PR 20-APR-1992; 92US-00870985.
PR 01-MAY-1992; 92US-00876941.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Tam AW, Reyes GR, Yarbough PO;
XX
DR WPI; 1997-558132/51.

DR N-PSDB; AAT96959.
XX Diagnosis of hepatitis E virus Burma and Mexico strain infection - by
PT immunoassay with hepatitis E virus peptide antigens.
XX
XX Disclosure; Col 39-42; 36pp; English.

XX A method has been developed for detecting hepatitis E virus (HEV)
CC antibodies (Ab). The method comprises: (a) reacting a serum sample with a
CC HEV peptide antigen; and (b) examining the peptide for the presence of
CC bound Ab, where the presence of bound Ab indicates the presence of HEV
CC Ab. The present presence represents the protein from the open reading
CC frame, ORF2, from HEV Burma strain. The method can be used to diagnose
CC infection with the enterically transmitted non-A/non-B viral hepatitis
CC agent HEV, specifically the HEV Burma and Mexico strains. Updated on 25-
CC MAR-2003 to correct PF field.) (Updated on 17-OCT-2003 to standardise OS
CC field)

XX Sequence 660 AA;

Alignment Scores:

Pred. No.:	0	Length:	660
Score:	660.00	Matches:	660
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAW35826 (1-660)

QY	1	ATCGCGCTCGGCTATTTGTCCTCTCATGTTTGGCTATGTCGCCGCGCCA	60
Db	1	MetArgProArgProIleLeuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20
QY	61	CCGCGCGCTCAGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCTTCCGCGGTGGT	120
Db	21	ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly	40
QY	121	TTCTGGGTCACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC	180
Db	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn	60
QY	181	CCCTTCGCCCGGATGTCACCGCTGCGCGCGCGCGCTGACCTCGTGTGCCAACCGCC	240
Db	61	ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla	80
QY	241	CGACCACTCGGCTCCGCTGGCGTGACCGCCAGCGCCCGCGCTGCTCAGCTCGT	300
Db	81	ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg	100
QY	301	AGACCTTACCACAGCTGGCGCGCGCTAACCGCGCTCGCTCCGCGCCATGACACCCCG	360
Db	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro	120
QY	361	CGAGTCCTGATGTCGACTCCCGCGCGCCATCTTCCGCGCGGAGTATACCTATCAACA	420
Db	121	ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr	140
QY	421	TCTCCCTTACCTCTTCGTTGGCCACCGGCACTAACTCGTCTTTTATGCGCGCTCTT	480
Db	141	SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu	160
QY	481	AGTCGCTTTTACCTTTCAGGACGGCACCAATACCATATATATGCGCACGGAAGCTTCT	540
Db	161	SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer	180
QY	541	AATATGCCAGTACCGGTTGCGGTGCCACAAATCGGTTACCGCGCGCTGTCCTCCCAAT	600
Db	181	AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn	200
QY	601	GCTGTGCGGCTTACGCACTCTCATCTCTTCTGCGCACAGACCAACCAACCGCGACG	660
Db	201	AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr	220

QY	661	TCGTTGATATGAATTCATTAACCTCGACGGATGTTCTGTTATTTTAGTCCAGCCGCGATA	720
Db	221	SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle	240
QY	721	GCCTCTGAGCTTGATCCCAAGTGAGCGCTACACTATCGTAAACCAAGCTGCGCTCC	780
Db	241	AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer	260
QY	781	GTCGAGACCTCTGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATTGCTTGCATA	840
Db	261	ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle	280
QY	841	CATGGCTCACTCGTAAATTCCTATCTAATACACCTATACCGGTGCGCTCGGCTGTTG	900
Db	281	HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu	300
QY	901	GACTTTGCCCTTGAGCTTGAGTTTCGAACCTTTACCCCGGTAAACCAATACGCGGTC	960
Db	301	AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal	320
QY	961	TCCCGTTATTCAGCACTCTCGCCACCGCTTCGTGCGGTGGGACGGAGCTGCCGAG	1020
Db	321	SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu	340
QY	1021	CTCACCAACGCGCTGCTACCCGCTTTATGAAGAGCTCTATTTTACTAGTACTAATGGT	1080
Db	341	LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly	360
QY	1081	GTCGGTAGATCGCGCGGATAGCCCTCACCTGTTTCACTTCTGCTGACCTCTGCTT	1140
Db	361	ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu	380
QY	1141	GCGGCGCTCGCAGCAATGATTTCGTGCGGTGGGCGGAGCTGTTACTCCCGTCCC	1200
Db	381	GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro	400
QY	1201	GTTGTCTCAGCAATGCGGAGCGGCTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG	1260
Db	401	ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln	420
QY	1261	CAGGATAAGGTATTGCAATCCCGATGACATTGACTCGGAGAACTCTGTTGGTTATT	1320
Db	421	GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle	440
QY	1321	CAGGATTATGATAACCAACATGAAGATCGCGCGAGCGCTTCTCAGCCCATCGCGC	1380
Db	441	GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg	460
QY	1381	CCTTTCTCTGCTTCGAGCTAATGATGCTTGGCTCTCTCTCACCGCTGCGGAGTAT	1440
Db	461	ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr	480
QY	1441	GACCACTCACTTATGGCTCTTCGACTGGCCCAAGTTATGTTCTGACTCTGTGACCTTG	1500
Db	481	AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu	500
QY	1501	GTTAATGTTGCGACCGCGCGCGGCTTGGCCGCTCGCTCGATTTGGACCAAGGTCA	1560
Db	501	ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr	520
QY	1561	CTTGAGCGTGGCCCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCGCG	1620
Db	521	LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro	540
QY	1621	CTCCGCGGTAAAGTCTCTTCTGGGAGCGAGGACCAACTAAAGCCGGTACCTTATAT	1680
Db	541	LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn	560
QY	1681	TATAACCACTGCTAGCGACCAACTGTTTGTGAGAATGCGCGCGGACCGCGGTGCT	1740
Db	561	TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla	580

QY 1741 ATTCCACTTACACACTAGCTGGGTGCTGGTCCGCTCTCCATTTCTTCGGGTTCGGCTT 1800
Db |||||
581 lIeSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATGTGTGAGTACCTTGAGTACCTCCCGCGCC 1860
Db |||||
601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGATGATTCTGCGCCAGAGTCCGCCCTTGCGCTTCAGGGCTGCGCTTC 1920
Db |||||
621 HisThrPheAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTGCGTCAAGTTCAGGCGCTTAAGATGAAGTGGTAAACTCGGAGTTG 1980
Db |||||
641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 5
AAW80197
ID AAW80197 standard; protein; 660 AA.
XX
AC AAW80197;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 23-DEC-1998 (first entry)
XX
DE Protein encoded by ORF2 of ET-NANB (HEV) Burma strain DNA sequence.
XX
KW Enterically transmitted non A non B hepatitis virus; ET-NANB;
KW Hepatitis E virus; HEV; Burma HEV isolate; vaccine; diagnostic probe.
XX
OS Non-A.
OS non-B hepatitis virus.
XX
PN US924649-A.
XX
PD 20-OCT-1998.
XX
PF 07-JUN-1995; 95US-00475807.
XX
PR 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-00336672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 25-JUL-1994; 94US-00279823.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Bradley DW, Reyes GR, Krawczynski KZ, Tam A, Fry KE, Yarbough PO;
XX
DR WPI; 1998-582599/49.
DR N-PSDB; AAV66321.
XX
PT Hepatitis E virus proteins - useful for diagnosis or vaccine production
PT the virus.
XX
PS Claim 6; Col 65-70; 47pp; English.
XX
CC AAW80196-98 are encoded by the genome of the Burma strain of enterically
CC transmitted non A non B hepatitis virus (ET-NANB) (hepatitis E virus
CC (HEV)). The specification describes an isolated protein which is
CC specifically immunoreactive with antibodies present in individuals
CC infected with HEV and encoded by a sequence contained in an open reading
CC frame (ORF) of an HEV genome. The genome has a sequence that is more than
CC 70% identical to the ORF1 sequence from Burma HEV isolate. The protein is
CC used as a vaccine and a diagnostic probe for ET-NANB. (Updated on 25-MAR-
CC 2003 to correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 660 AA;

Alignment Scores:
Pred. No.: 0 Length: 660

Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAW80197 (1-660)

QY 1 ATGCGCCCTCGGCTCATTTTGTGTGCTGCTCTCATGTTTGTGCTATGCTGCGCGGCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCCCGGTCAGCCCTCTGCGCGCGCTGCGCGCGCGCAGCGGGGTTCGCGCGGTGT 120
Db |||||
21 ProProGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlySerGlyGlyGly 40
QY 121 TTCGGGTGACCCGGTGTATTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db |||||
41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCGCCGATGTACCGCTGCGCGCGCGGCTGGACCTCGTGTTCGCAACCCGCC 240
Db |||||
61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTTGGCTGACCAAGCGCCAGCGCCCGCTTCCTCAGCTCGT 300
Db |||||
81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACTTACACAGCTGGCGCGCGCTTAACCGCGTGTCTCGCGCCCATGACACCCCG 360
Db |||||
101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTGACTCCCGCGCGCCCATCTTTCGCGCGCGCAGTATTAACCTACA 420
Db |||||
121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCTGGCCACCGCGCACTAACCTGTTCTTTATGCCCGCCCTCT 480
Db |||||
141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTACCCCTTCAGSACGGCACCAATACCCATATAATGCCACGGAAGTTCT 540
Db |||||
161 SerProLeuLeuProLeuGluAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTCGCGCGCCACCAATCGTTACCGCGCGCTGCTCCCAAT 600
Db |||||
181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGGCGGTTACGCCATCTCCCATCTCATCTCTGCGCCACAGACCACACCCCGCAG 660
Db |||||
201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATGATCAATTCAATTAACCTCGACGAGTGTTCGTATTATTAGTCCAGCCCGCAT 720
Db |||||
221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTGTGATCCCAAGTGAAGCGCTACACTATCGTAACAGGCTGGCGCTCC 780
Db |||||
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACTCTCGGGTGGCTGAGGAGGAGGTACTCTGCTGTTGTTTATGTTTGCAT 840
Db |||||
261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCACTCGTAAATTCCTATACTAATACCCCTATACCGGTGCGCTCGGGCTGTG 900
Db |||||
281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTCGCTTGAGCTTGAGTTTCGAACTTACCCCGCGGTAAACCAATACGCGGCTC 960
Db |||||
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTGCGGGTGGGAGGACTGCCGAG 1020

Db 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACCTACACAGCTGGGGCGCGCGCTAAACCGGGTGGCTCCGGCCCATACACCCCG 360
Db 101 ArgProThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTGCAGCTCCCGCGCGGCATCTTGGCGCGGCAGTATAACTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTCCGTGGCCACCGCACATAACCTGGTCTTTATGCCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGGCTTTACCCCTTCAGACCGCACCATACCATATAATGCGCCACGGAAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTCCCGGTGCCCAATCCGTTACCGCGCTGCTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGGCGTTACGCCATCTCCATCTCATCTCGGCCACAGACACACACCCCGCAG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATACCTCGACGGATGTCGTATTTAGTCAGCCCGGCATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTGTGATCCCAAGTCAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACTCTGGGGTGGCTGAGGAGAGGTACTCTGTCGTCTGTATGCTTTGCCATA 840
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCACTCGTAATTCCTATACTATAACCTATACCGGTACCGGTGCCCTCGGCTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGGCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACGGGGTTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCCAGCACTGCTCGCCACCGCCTTCGTCGGGTGCGGACGGGACTGCCGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCCACCGGCTGCTACCGCTTATGAGGACCTCTATTTACTAGTACTAATGGT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGCTGAGATCGCGCGGGTAGCCCTCACCCCTGTTCAACCTTGCTGACACTCTGCTT 1140
Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTTCCGACAGAAATTTGTCGTCGGCTGGTGGCCAGCTGTTCTACTCCCGTCCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGCTCAGCAATGGCGAGCGGACTGTAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGGTTATT 1320
Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGlySerArgValIle 440
QY 1321 CAGGATTATGATACCAACATGAACAGATCGCGCAGCCCTTCTCCAGCCCATCGCGC 1380

Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTTCTCTGCTTCGAGCTAAATGATGCTGTTTGGCTCTCTCTCACCGCTGCCGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCAGTCCACTTATGGCTCTTCGACTGGCCAGCTTTATGTTTCTGACTCTCTGTGACCTTG 1500
Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGGACACCGCGCGCAGCGCTGGTCCCGTGGCTCGATTGACACCAAGGTACA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTTGACGGTCCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTGCTCGCG 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCCCGCGTAAAGCTCTCTTTCTGGGAGGCAGGCACAACTAAAGCGGGTACCTTTATAT 1680
Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACACTGTAGCGACCAACTGCTTGTGAGAAATGCCCGCGGCACCGGTGCGT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGTCGCTCCGCTCCATTTCTCGGTTCCGCTT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGGACTACCTCCCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTGATGATTTCTGCCAGTGCAGTCCGCCCTTGGCCTTCAGGGCTGCGCTTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTGAGCTTCAGCGCTTAAGATGAAGTGGTAAACCTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 7
AAB24120
ID AAB24120 standard; protein; 660 AA.
XX AAB24120;
AC AAB24120;
DT 29-JAN-2001 (first entry)
XX
DE Hepatitis E virus Burma strain ORF2 protein sequence SEQ ID NO:8.
XX
KW Hepatitis E virus; HEV; non-A/non-B viral hepatitis; viral protein; immunological; diagnosis; hepatitis; infection; identification; detection; immunoreactive; hepatotropic; antiinflammatory; virucide; vaccine; antiviral; antigenic; antibody; antigen.
XX
OS Hepatitis E virus.
XX
PN US6120988-A.
XX
PD 19-SEP-2000.
XX
PF 07-JUN-1995; 95US-00478507.
XX
PR 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-00336672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 05-APR-1991; 91US-00681078.
PR 25-JUL-1994; 94US-00279823.
XX

PA (GENE-) GENELABS TECHNOLOGIES INC.
XX (USGO) US GOVERNMENT.
XX Yarbough PO, Krawczynski KZ, Fry KE, Bradley DW, Tam A, Reyes GR;
PI N-PSDB; AAA99259.
XX WPI; 2000-593712/56.
DR N-PSDB; AAA99259.
XX
PT Identifying recombinant antigen immunoreactive with antibody induced by
PT hepatitis E virus (HEV), for detecting HEV infection, comprises
PT immunoreacting a polypeptide from an HEV genome with an HEV-positive
PT antiserum.
XX
PS Claim 6; Col 63-68; 46pp; English.
XX
CC The present invention describes a method for identifying a recombinant
CC antigen immunoreactive with a hepatitis E virus (HEV)-induced antibody.
CC The method comprises producing a polypeptide derived from an HEV genome,
CC immunoreacting the polypeptide with an HEV-positive antiserum and
CC selecting the polypeptide as a recombinant antigen if the polypeptide
CC reacts with the HEV-positive antiserum. The method is useful for
CC identifying recombinant antigen immunoreactive with antibody induced by
CC HEV. The enterically transmitted non-A/non-B hepatitis (ET-NANB (also
CC known as HEV))-specific fragments are useful for identifying ET-NANB-
CC derived cDNAs, which contain additional sequence information, as primers
CC for detecting ET-NANB viral genomic material in a patient sample, for the
CC synthesis of polypeptides for use in immunoassays, and for identifying
CC similar antigenic regions encoded by related viral strains, e.g. Burmese
CC strain. The antigens are especially useful in the preparation of vaccine
CC against ET-NANB infection. These antigens may further be used to prepare
CC antibodies to ET-NANB virus particles for use directly as antiviral
CC agents, and to produce antiserum designed for pre- or post-exposure
CC prophylaxis. The present sequence represents a specifically claimed HEV
CC Burma strain protein sequence for use in the present invention
XX
SQ Sequence 660 AA;

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAB24120 (1-660)

QY 1 ATGGCGCCTCGGCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCGCGGCCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20

QY 61 CCGCGCGGTCCAGCGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGGTCCGCGGTGGT 120
DB 21 ProProGlyGlnProSerGlyArgArgGlyArgGlyArgSerGlyGlySerGlyGly 40

QY 121 TTCTGGGGTGACCGGGTGTATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60

QY 181 CCCTTGCCCGCCGATGTCACCGTGGCGCGGGGTGGACCTCGTGTGCGCCAAACCGGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80

QY 241 CGACCACTCGGCTCCGCTGCGGTGACCGCGCCAGCGCGCGCGCGGTGCTCAGCTCGT 300
DB 81 ArgProLeuGlySerAlaIleProArgGlnAlaGlnArgProAlaValAlaSerArgArg 100

QY 301 AGACCTACCAAGTGGGGCGCGCGCTAACCGCGTGGCTCGCGCGCCATGACACCCCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuLeuThrAlaValAlaProAlaHisAspThrPro 120

QY 361 CAGTGGCTGATGTCGACTCCGCGCGCGCATCTTGGCGCGCGAGTATACCTATCAACA 420
DB 481 AspGlnSerThrTyrGlySerThrGlyProValTyrValSerAspSerValThrLeu 500

Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCCGTGGCCAGCGCACAACTGCTGTTCTTTATCGCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTTACCTCTCAGGACCGCACCAATACCATATTAATGGCCAGGAAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTCCTGCGCACCAATCCGTTACCGCGCTGGTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaAlaAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCCGGGTTACGCCATCTCATCTCATTTCTGGCCACAGACCAACCCCGACG 660
Db 201 AlaValGlyGlyPyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATAACCTCGACGATGTTCTGTATTTTAGTCCAGCCCGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATCGTAAACAGGCTGGCGTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGCTCTGCTTGTCTTGCATA 840
Db 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTCGTAAATTCCTATCTATACACCTATACCGGTGCCCTCGCGCTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACTTACCCCGGTAAACCAATACCGGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCCAGCACTGCTGCCACCGCTTCTGCGGTGCGGAGCGGAGCTGCCGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCAACCGCTGCTACCCGCTTTATGAAGAGCACTCTATTTTACTAGTACTATGGT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGGTGAATCGCGCGGAGATAGCCCTCACCTGTTTCAACCTGCTGTCGACTGCTCT 1140
Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuLeuThrPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCCGACAGAAATTTGATTCTGCGGTGGTGGCCAGCTGTTCTACTCCCGTCCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCCCAATCGCGAGCCGCACTGTTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGGTATTGCAATCCCGATGACATTCAGCTCGGAGAACTCGTGTGGTTATT 1320
Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValAlaIle 440
QY 1321 CAGGATTATGATAACCAACATGAACAGATCGCGCCAGCGCTTCTCAGCCCATCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGCTCTCGAGCTAATGATGCTTTGCTCTCTCTCAGCGCTGCGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCCACTATGGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGAGCCTTG 1500
Db 481 AspGlnSerThrTyrGlySerThrGlyProValTyrValSerAspSerValThrLeu 500

QY	1501	GTAAATGTTGCGACCGCGCGAGCGCGTTCGCCGGTCTCGATTGGACCAAGGTCA	1560
Db	501	ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrylsValThr	520
QY	1561	CTTGACGGTCGCCCTCTCCACATCCAGAGTACTCGAAGACCTTCTTGTCTCGCG	1620
Db	521	LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuPro	540
QY	1621	CTCCGGGTAAAGTCTCTTCTGGGAGGAGGACACAACTAAAGCCGGTACCTTATAAT	1680
Db	541	LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrLysAlaGlyTyrProTyrAsn	560
QY	1681	TATAACACCACTGCTAGCGACCAACTGCTGTGCGAATGCGCGCGCACCGGGTCGCT	1740
Db	561	TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaGlyHisArgValAla	580
QY	1741	ATTTCACCTACACCACTAGCTGGTGTGTCGCTCCGCTCTCCATTTCTCGGTTGCCGT	1800
Db	581	IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal	600
QY	1801	TTAGCCCCCACTCTCGCGTAGCTGCTTGGAGTACCTTGGACTACCTGCGCGCGCC	1860
Db	601	LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla	620
QY	1861	CATACCTTTTGATGATTCTGCGCCAGAGTCCCGCCCTTGGCTTCAGGGCTGCGCTTC	1920
Db	621	HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe	640
QY	1921	CAGTCTACTGTCGCTGAGCTTCAGCGCCTTAAGATGAAGTGGTAAAACTCGGAGTTG	1980
Db	641	GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu	660
RESULT 8			
AAB62523			
ID	AAB62523 standard; protein; 660 AA.		
AC	AAB62523;		
DT	23-JUL-2001 (first entry)		
DE	HEV-Burma strain viral protein ORF2.		
KW	HEV; enterically transmitted nonA/nonB viral hepatitis agent; viral;		
KW	ET-NANB; infection; vaccine.		
OS	Hepatitis E virus.		
XX	US6229005-B1.		
XX	08-MAY-2001.		
XX	03-AUG-1998; 98US-00128275.		
XX	17-JUN-1988; 88US-00208997.		
PR	11-APR-1989; 89US-00336672.		
PR	16-JUN-1989; 89US-00367486.		
PR	13-OCT-1989; 89US-00420921.		
PR	05-APR-1990; 90US-00505888.		
PR	05-APR-1991; 91US-00681078.		
PR	25-JUL-1994; 94US-00279823.		
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.		
PA	(GENE-) GENELABS TECHNOLOGIES INC.		
XX	Reyes GR, Yarbrough PO, Bradley DW, Krawczynski KZ, Tam AW;		
PI	Fry KE;		
XX	WPI; 2001-342705/36.		
DR	N-PSDB; AAF83495.		
XX	New DNA sequences of enterically transmitted non-A/non-B (ET-NANB)		
PT	hepatitis viral agent, useful in diagnosing infection by an enterically		

PT	transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.			
XX				
PS	Disclosure; Col 63-68; 45pp; English.			
XX				
CC	The invention relates to an isolated DNA comprising the genome of an			
CC	enterically transmitted nonA/nonB (ET-NANB) viral hepatitis agent (also			
CC	referred as HEV). The DNA sequences or their fragments are useful in			
CC	preparing ET-NANB viral proteins and as probes for virus detection. These			
CC	are particularly useful in diagnosing infection by an enterically			
CC	transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.			
CC	The present sequence represents a ORF2 protein encoded by a ET-NANB viral			
CC	DNA sequence from HEV-Burma strain			
XX				
SQ	Sequence 660 AA;			
Alignment Scores:				
Pred. No.:	0	Length:	660	
Score:	660.00	Matches:	660	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	4	Gaps:	0	
US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAB62523 (1-660)				
QY	1	ATGCGCCTCGGCTATTTTGTGCTCCTCATGTTTTTGCTATGCTGCCGCGCCA	60	
Db	1	MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20	
QY	61	CGCGCGCTCAGCGCTCTGCGCGCGCTGCTGCGCGCGCGCGGTTCCGCGCGTGT	120	
Db	21	ProProGlyGlnProSerGlyArgArgGlyArgGlySerGlyGlySerGlyGlyGly	40	
QY	121	TTCTGGGTGACCGGTTGATTCAGCCCTTCGCAATCCCTATATTATTCATCAACAAC	180	
Db	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn	60	
QY	181	CCCTTCGCGCCCGATGTCACCGCTCGCGCGCGGCTGGACCTCGTTCGCGCAACCGCGC	240	
Db	61	ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla	80	
QY	241	CGACCACTCGGCTCGGCTTGGCGTGACAGGCGCCAGCGCGCGCTTCGCTCAGCTCGT	300	
Db	81	ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg	100	
QY	301	AGACCTACACAGCTGGGCGCGCGCTAACCGCGGTGCTCGCGCCCATGACACCGCG	360	
Db	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro	120	
QY	361	CCAGTGCCTGATGTCGACTCCCGCGCGCGCATCTTGGCGCGCGCAGTATACCTATCAACA	420	
Db	121	ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr	140	
QY	421	TCTCCCTTACCTCTTCCGTGGCCACCGGCATAACCTGGTTCTTTATCCCGCGCCCTCTT	480	
Db	141	SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu	160	
QY	481	AGTCGCGCTTTACCCCTTCAGACCGGCACCAATATATATATGCGCCACGGAAGCTTCT	540	
Db	161	SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer	180	
QY	541	AATATATGCCAGTACCGGTTGCCGTGCCCAATCCGTACCGCGCGCTGCTGCCCAT	600	
Db	181	AsnTyrAlaGlnTyrArgValAlaAlaAlaThrIleArgTyrArgProLeuValProAsn	200	
QY	601	GCTGTGCGCGGTTACGCCATCTCCATCTCATTTCTGGCCACAGACCAACACCGCGAG	660	
Db	201	AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr	220	
QY	661	TCCGTTGATATGAATTCAATAACCTCGACGGATGTCGTATTTTATAGTCCAGCCCGCAT	720	
Db	221	SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle	240	

QY 721 GCCTCTGAGCTTGATCCCAAGTGAGCGCTACACTATCTGTAACCAAGGCTGCGCTCC 780
Db |||||
241 AlaserGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrPargSer 260
QY 781 GTCGAGACCTCTCGGGTGGCTGAGAGGAGGCTACCTCTCGGTCTTGTATGCTTTGCATA 840
Db |||||
261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTCGTAATCTCTATACATAACACCTATACCGGTGCGGCTGCTTG 900
Db |||||
281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTCAGCTTGAAGTTTCGCAACTTACCCCGGTAAACCAATACGCGGTC 960
Db |||||
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCACGACTGCTCGCCACCGCTTTCGTCGCGGTGCGGACGCGGCTGCGAG 1020
Db |||||
321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACACACGCTGCTACCGCTTATGAAGACCTCTATTTACTAGTACTACTATGCT 1080
Db |||||
341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGGTCGAGATCGCGCGGATAGCCCTCACCCCTGTTCACCTTGTCTGACACTCTGCTT 1140
Db |||||
361 ValGlyGluLeuGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTGCGACAGAAATTGATTTCTCGCGGTGCGGCGAGCTTCTACTCCCGTCCC 1200
Db |||||
381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATGCGGACCGACTTAAAGTTGTATACATCTGAGAGATGCTCAG 1260
Db |||||
401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGATTTGCAATCCCCCATGACATTCGACCTCGGAGAAATCTCGTGTGTTATT 1320
Db |||||
421 GlnAspLysGlyIleAlaIleProHisIleAspLeuGlyGluSerArgValValIle 440
QY 1321 CAGGATATGATAACCAACATGAACAAGATCGGCGGACGCTTCTCAGCGCCCATCGCGC 1380
Db |||||
441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTCTCTCTCTCCAGCTTAAGTATGCTTGGCTCTCTCTCAGCGTCCGCGACTAT 1440
Db |||||
461 ProPheSerValLeuArgAlaAsnAspValLeuThrPheSerLeuLeuAlaAlaGluTyr 480
QY 1441 GACGAGTCCACTTATGGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db |||||
481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGGACCGGCGGACGCGTTCGCGGTGCGCTCGATTGGACCAAGGTACA 1560
Db |||||
501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrThrLysValThr 520
QY 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCGCTCCG 1620
Db |||||
521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAGCTCTCTTTCTGGGAGCGAGGACAACTAAAGCGGGTACCCCTATAAT 1680
Db |||||
541 LeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACCACTAGTAGGACCACTGCTTGTGAGAAATCGCGCGGCGGACCGGTGCT 1740
Db |||||
561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACCACTAGCTGGGTGCTGCTCGCTCTCCATTTCTGCGGTGGCGCTT 1800
Db |||||
581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGTTGAGGATACCTTGGACTACCTCGCGCGGCC 1860

Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGAATTTCTGCCAGAGTGCAGAGTCCCGCCCTTGGCTTCAGGCTCGCTTC 1920
Db |||||
621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGCTCTACTGCTCGCTCAGCTTTCAGCGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 1980
Db |||||
641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660
RESULT 9
AAOI5699
ID AAO15699 standard; protein; 660 AA.
XX AAO15699;
XX 08-NOV-2002 (first entry)
XX Hepatitis E virus (Burma strain) ORF2 protein.
XX HEV; enterically transmitted nonA/nonB hepatitis viral agent;
KW Burma strain; bile; ORF2.
XX Hepatitis E virus.
XX US6379891-B1.
XX 30-APR-2002.
XX 19-APR-2000; 2000US-00553427.
XX 17-JUN-1988; 88US-00208997.
XX 11-APR-1989; 89US-00336672.
XX 16-JUN-1989; 89US-00367486.
XX 13-OCT-1989; 89US-00420921.
XX 05-APR-1990; 90US-00505888.
XX 15-APR-1991; 91US-00681078.
XX 25-JUL-1994; 94US-00279823.
XX 07-JUN-1995; 95US-00478507.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam A, Fry KE;
XX WPI; 2002-517277/55.
XX N-PSDB; AAL50386.
XX Detecting the presence of enterically transmitted nonA/nonB hepatitis viral (HEV) agents in bile samples from infected humans and monkeys using polymerase chain reaction.
XX Disclosure; Col 19-24; 61pp; English.
XX The invention comprises a method for detecting the presence of enterically transmitted nonA/nonB hepatitis viral (HEV) agents in a sample and isolating HEV agents or nucleic acid fragments produced by the agent. The method utilises PCR - using bile from a human or cynomolgus monkey actively infected with HEV as a source of the agent. The method of the invention is used for detecting the presence of a viral agent in a sample of cultured cells infected with the agent and isolating enterically transmitted nonA/nonB HEV agents or nucleic acid fragments produced by the agent. The present amino acid sequence represents the protein encoded by open reading frame 2 (ORF2) of a DNA sequence isolated from a Burmese strain of the Hepatitis E virus
XX Sequence 660 AA;
SQ
Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00%			Mismatches: 0			
Query Match: 100.00%			Indels: 0			
DB: 5			Gaps: 0			
US-09-851-410A-6_COPY_5147_7129 (1-1983) x AA015699 (1-660)						
QY	1	ATGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTTCCTATGCTGTCGCGGCCA	60	QY	1021	CTCACCACACGGCTCTACCGCTTTATCAAGGACCTCTATTTTACTAGTACTAATGGT
DB	1	MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20	DB	341	LeuThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly
QY	61	CGCGCGGTGAGCGCTGCTCGCGCGTGTGTGGCGCGCGCAGCGCGGTTCGCGGGTGT	120	QY	1081	GTCCGTGAGATCGCGCGGGATAGCCCTCACCCCTGTTCAACCTTCTGACACTCTGCTT
DB	21	ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly	40	DB	361	ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu
QY	121	TTCTGGGGTCAACGGGTGTATCTTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC	180	QY	1141	GGCGGCTCGCAGACAAATTGATTTCTGCGGTGTGGCCAGCTGTTCTACTCCCGTCCC
DB	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn	60	DB	381	GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro
QY	181	CCCTTCGCCCGGATGTACCGTGTGCGCGCGGGCTGGACCTGCTGTTCGCCAACCGCC	240	QY	1201	GTTCGTCTCAGCCCAATGCGGAGCGGAGCTGTTAAGTTTGTATATCATCTCTAGAGAATGCTCAG
DB	61	ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla	80	DB	401	ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln
QY	241	CGAACCTCGCTCGCTTCGCGTGACAGGCCCGCGCGCGCTGCCTCAGCTCGT	300	QY	1261	CAGGATAAGGGTATTGCAATCCCGCATGACATTCGATCGGAGAAATCTCGTGTGGTTATT
DB	81	ArgProLeuGlySerAlaIleProArgGlnAlaGlnArgProAlaValAlaSerArgArg	100	DB	421	GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle
QY	301	AGACCTACACAGCTGGGCGCGCGCTAAACCGCGTGTCTCCGCGCCATGACACCCCG	360	QY	1321	CAGGATTATGATAACCAACATGAACAAGATCGCGCGAGCGCTTCTCCAGCCCATCGCGC
DB	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro	120	DB	441	GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg
QY	361	CCAGTGCCTGATGTGACTCCCGCGCGCCATCTTTCGCGCGCAGTATAACCTATCAACA	420	QY	1381	CTTTCTCTGTCTTCGAGCTAAATGATGTCTTTGGCTCTCTCTCACCGCTCCCGAGTAT
DB	121	ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr	140	DB	461	ProPheSerValLeuArgAlaAsnAspValLeuThrLeuSerLeuThrAlaAlaGluTyr
QY	421	TCTCCCTTACTCTTCGCTGGCCACCGGCACCTGCTTCTTATGCGCGCCCTCTT	480	QY	1441	GACCACTGCATTTATGGCTCTTCGACTGGCCAGTTATGTTTCTGACTCTGTGACCTTG
DB	141	SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu	160	DB	481	AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu
QY	481	AGTCGCGTTTACCCCTTCAGGACGGCACCAATACCATATATGCGCCAGGAAGCTTCT	540	QY	1501	GTTAATGTTGCGACCGCGCGCGCGTGTGCGCGTGTGCGCTCGATGAGACCAAGGTACA
DB	161	SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer	180	DB	501	ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr
QY	541	AATTATGCCAGTACCGGGTTGCGCGTGCCACAATCCGTTACCGCGCGTGTGCCCAAT	600	QY	1561	CTTGACGGTTCGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGCG
DB	181	AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn	200	DB	521	LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro
QY	601	GCTGTGCGGGTTACGCATCTCCATCTCATCTGCGCCACAGACCAACACCGCGAGC	660	QY	1621	CTCCGCGGTAGCTCTCTTTCTGGGAGCGAGGCACAACTAAAGCCGGGTACCTTTATAAT
DB	201	AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr	220	DB	541	LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn
QY	661	TCCGTTGATATGAATTCATAAATCCTCGACGATGTTCTGTTATTTAGTCCAGCGCGCAT	720	QY	1681	TATAACACACTGTAGCGACCAACTGCTTGTGAGAAATGCGCGCGGACCGGGTCGCT
DB	221	SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle	240	DB	561	TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla
QY	721	GCTCTGAGCTGTGATCCCAAGTAGCGCTACACTATCGTAACCAAGCTGCGCGTCC	780	QY	1741	ATTTCACCTTACACCACTAGCCTGCGTGTGCTCCGCTCTCCATTTCTGCGGTGCGGTT
DB	241	AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer	260	DB	581	IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal
QY	781	GTCGAGACCTCTGGGTGGCTGAGGAGGCTACCTCTGCTGTGTTATGCTTGCATA	840	QY	1801	TTAGCCCCCACTCTCGCTAGCAATTGCTTGAGGATACCTTGACTACCTGCGCGCGCC
DB	261	ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle	280	DB	601	LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla
QY	841	CATGCTCACTCGTAAATTCCTATACATAACCTATACCGGTGCCCTCGGCTGTG	900	QY	1861	CATACCTTTGATGATTTCTGCGCAGAGTCCGCCCTTGGCCTTCAGGGCTGCGCTTTC
DB	281	HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu	300	DB	621	HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe
QY	901	GACTTTGCCCTTGAGCTTCAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGGTC	960	QY	1921	CAGTCTACTGTCTGCTGAGCTTCAGCGCCTTAAAGTAGAAGGTGGGTAAAACTCGGGAGTTG
DB	301	AspPheAlaLeuLeuLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal	320	DB	641	GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu
QY	961	TCCCGTTATCCAGCACTGTCCGACCGCCTTCGTGCGGGTCCGAGACGGGACTGCCGAG	1020	RESULT 10		
DB	321	SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu	340	ADD24376		
				ID	ADD24376	standard; protein; 660 AA.
				XX	XX	
				AC	ADD24376;	
				DT	15-JAN-2004	(first entry)
				XX	XX	

QY 1321 CAGGATTATGATAACCAACATGAAACAGATCGCGAGCGCTCTCCAGCCCACTCGCGC 1380
Db |||||
441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTCTCTGCTCTCGAGCTAATGATGCTTGGCTCTCTCTCACCGCTGCCAGTAT 1440
Db |||||
461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTGCTTATGGCTCTCGACTGCGCCAGTTTATGTTCTGACTCTGTGACCTTG 1500
Db |||||
481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTATGTTGGACCGCGCGCAGCGCTGTCGCCGTGCTCGATTGGACCAAGTCA 1560
Db |||||
501 ValAsnValAlaThrGlyAlaGlnAlaValAlaAArgSerLeuAspTrpThrIysValThr 520
QY 1561 CTTGACGTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTCGCG 1620
Db |||||
521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPheValLeuPro 540
QY 1621 CTCGCGGTAAAGCTCTCTTCTGGGAGCGAGCACAACCTAAAGCGGGTACCTTATAT 1680
Db |||||
541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAAATGCGCGGCGCACCGGTCGCT 1740
Db |||||
561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTCCACTTACACACTAGCTGGGTGCTGCTCCGCTCTCAATTCGTGGGTGCCGTT 1800
Db |||||
581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTGCTTGAGGATACCTTGACTACCTGCCCGCGC 1860
Db |||||
601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACCTTTGATGATTCTGCGCCAGAGTCCGCCCTTGGCTTTCAGGGCTGCGCTTTC 1920
Db |||||
621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGAGCTTACGCGCTTAAGAGTGAAGGTGGTAAACTCGGAGTTG 1980
Db |||||
641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 11

AAW76369 standard; protein; 660 AA.

AAW76369;

03-DEC-1998 (first entry)

Hepatitis E virus hollow particle protein #2.

Hollow particle protein; virus; antibody; detection; immunoassay; infection.

Hepatitis virus.

JP10234383-A.

08-SEP-1998.

28-FEB-1997; 97JP-00062445.

28-FEB-1997; 97JP-00062445.

(DENK-) DENKA SEIKEN KK.

(KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.

WPI; 1998-535037/46.

N-PSDB; AAV61688.

PT Hepatitis E virus hollow particle poly(peptide(s) and nucleic acids
PT encoding it - useful for more accurate detection of HEV in samples, using
PT immuno-assays and nucleic acid hybridisation.

XX Claim 13; Page 24-26; 29pp; Japanese.

XX This sequence represents a Hepatitis E viral hollow particle protein.
CC This polypeptide can be used to raise antibodies to detect HEV infection
CC in samples, e.g. by immuno-assay based techniques, and the nucleic acid
CC can be used for the same in nucleic acid hybridisation assays. The
CC polypeptides and nucleic acids allow more accurate detection of HEV than
CC previously possible

XX Sequence 660 AA;

Alignment Scores:

Pred. No.: 0 Length: 660
Score: 559.00 Matches: 659
Percent Similarity: 99.85% Conservative: 0
Best Local Similarity: 99.85% Mismatches: 1
Query Match: 84.70% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAW76369 (1-660)

QY 1 ATGGCGCCTCGGCTATTTTGTGCTGCTCTCATGTTTTCCTATGCTGCGCGGCCA 60

Db 1 MetArgProArgProIleLeuLeuLeuLeuLeuMetPheLeuProMeLeuProAlaPro 20

QY 61 CCGCCCGCTCAGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGTTCGCGCGGTGT 120

Db 21 ProProGlyGlnProSerGlyArgArgArgGlyArgArgSerGlyGlyGlyGly 40

QY 121 TTCTGGGTGACCGGTGATCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC 180

Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60

QY 181 CCCTTGGCGCCCGATGTCCACCGCTGCGCGCGGTGGACCTCGTGTTCGCCCAACCCGCC 240

Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80

QY 241 CGACCACTCGGCTCGCTTGGCTGGCGACCGACCGCGCGCGCGCGCTTGCCTCAGCTGT 300

Db 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100

QY 301 AGACTACACAGCTGGCGCGCGCGCTAACCGCGGTGCTCGCGCGCGCGCGCGCGCG 360

Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120

QY 361 CCAGTGCCTGATGTGACTCCCGCGCGCGCGCTTGGCGCGCGCGCTATTAACCTATCAACA 420

Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140

QY 421 TCTCCCTTACTCTCTCGTGGCCACCGCGCACTAACCTGGTTCCTTTATGCGCGCCCTCT 480

Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160

QY 481 AGTCGCTTTTACCCCTTCAGGACGGCAACCAATCCCATATAATGCGCACGAGACTTCT 540

Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180

QY 541 AATTATGCCAGTACCGGGTTGCCGCTGCCCAATCCGTTACCGCGCGCGCTGCCCAAT 600

Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200

QY 601 GCTGTGCGCGGTATACGCCATCTCCATCTCTTCTGGCCACAGACCAACCCACCGCGACG 660

Db 201 AlaValGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220

QY 661 TCCGTTGATGATTAATTAACCTCGACGGAATGTTGATATTTAGTCCACCGCGCGATA 720

Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240

QY 721 GCCTCTGAGCTTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
Db |||||
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGTGGCTGAGAGAGAGGCTACCTCTGGTCTTGTATTGCTTGCATA 840
Db |||||
261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTCGTAATTCCTATATACTATACACCTATACCGGTCCCTCGGGCTGTG 900
Db |||||
281 HasGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTCGAGTTTCGCAACTTACCCCGGTAAACCAATACGCGGTC 960
Db |||||
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGTTATTCAGACACTGCTGCCACCGGCTTCGTCGGGTGGCGGACGGAGCTCCGAG 1020
Db |||||
321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACGAGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGTT 1080
Db |||||
341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGCTGAGATCGGCGCGGATAGCCCTCACCTCTTCAACCTTCGCTGACACTCTGCTT 1140
Db |||||
361 ValGlyGluIleGlyArgGlyAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGGCTCGGACAGAAATTGTTGCTGGTGGTGGCGAGCTGTCTACTCCCGTCCC 1200
Db |||||
381 GlyGlyLeuProThrGluLeuLeuSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATCGCGAGCGACTGTTAGCTTGTATACATCTGTAGAGAACTCTAG 1260
Db |||||
401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGGTATTGCAATCCGCGATGACATTGACCTCGGAGAACTCTGTGGGTATT 1320
Db |||||
421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
QY 1321 CAGGATTATGATAACCAATGACAGATCGCGCCGAGCCCTCTCCAGCCCATCGGC 1380
Db |||||
441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGCTCTCGAGCTAATGATGTGCTTGGCTCTCTCACCCTGCGCGAGTAT 1440
Db |||||
461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCCACTTATGGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db |||||
481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTATGTTGCGACCGCGCGAGCGCTTGGCGGTCTGCTCGATTGGACCAAGTCA 1560
Db |||||
501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTTGACGGTGGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTCTTTGTCCTCGCG 1620
Db |||||
521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPheValLeuPro 540
QY 1621 CTCGCGGGTAAGCTCTTTTGGGAGGAGGAGGACCAACTAAAGCCGGTACCTTTATAAT 1680
Db |||||
541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATACACCACTGCTAGACACCAACTGCTTGTGAGAATGCCGCGGACCGGTCGCT 1740
Db |||||
561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCACCTTACCACTAGCTCGGTGCTGTGTCCTGCTTCCATTTCTCGGGTTCGGTT 1800
Db |||||
581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTCGGCTAGCATTTGCTTGGAGTACCTTGGACTACCTCGCCCGGCC 1860

Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTATGATGTTCTGCCAGAGTCCGCCCTTGGCTTCAGGCTCGGCTTTC 1920
Db |||||
621 HisThrPheAspAspPheCysProGluCysArgProLeuGluGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTCAGCTTCAGCGCCTTAAGATGAAGTGGGTAAAACTCGGAGTTG 1980
Db |||||
641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 12

AAW71210

ID AAW71210 standard; protein; 660 AA.

XX AAW71210;

XX 25-MAR-2003 (revised)

DT 30-OCT-1998 (first entry)

XX Protein encoded by ORF 2 of the Burmese isolate of ET-NANB.

XX Enterically transmitted nonA/nonB hepatitis virus; identification; HEV;

KW ET-NANB; detection; vaccine.

XX Hepatitis virus.

XX Key Location/Qualifiers

FH Misc-difference 496

FT /note= "not specified"

XX US5789559-A.

XX 04-AUG-1998.

XX 25-JUL-1994; 94US-00279823.

XX 17-JUN-1988; 88US-00208997.

PR 11-APR-1989; 89US-00336672.

PR 19-JUN-1989; 89US-00367486.

PR 13-OCT-1989; 89US-00420921.

PR 05-APR-1990; 90US-00505888.

PR 05-APR-1991; 91US-00681078.

XX (GENE-) GENELABS TECHNOLOGIES INC.

XX Bradley DW, Reyes GR, Krawczynski KZ, Tam A, Fry KE, Yarbough PO;

XX WPI; 1998-446186/38.

XX Hepatitis E virus DNA - useful for e.g. virus detection and viral protein production.

XX Disclosure; Col 63-68; 45pp; English.

XX AAW71209-11 represent the proteins encoded by the open reading frames (ORFs) of the DNA sequence of the Burmese isolate of an enterically transmitted nonA/nonB viral hepatitis agent (ET-NANB). The nucleic acid sequence may be used for identifying and sequencing the entire viral agent (also referred to as HEV), detecting ET-NANB in infected samples, e.g. by specific amplification of virus-derived DNA sequences and for producing recombinant viral proteins for use in vaccines. (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 660 AA;

Alignment Scores:

Pred. No.:	0	Length:	660
Score:	559.00	Matches:	659
Percent Similarity:	99.85%	Conservative:	0
Best Local Similarity:	99.85%	Mismatches:	1
Query Match:	84.70%	Indels:	0
DB:	2	Gaps:	0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAW71210 (1-660)

QY 1 ATGGCGCCTCGGCTATTTGCTGCTCTCTATGCTTTTGGCTATGCTGCGCGCGCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20

QY 61 CCGCGCGGTTCAGCGGTCTGGCGCGCGCTGCTGGCGCGCGAGCGCGGTTCGGCGGTGGT 120
DB 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly 40

QY 121 TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60

QY 181 CCCTTCGCCGCCGATGTACCGGTGCGCGCGGGGCTGACCTGTGTTCGCGCAACCGGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80

QY 241 CGACCACTTCGGCTCGGCTGGCGGTGACCGCGCGCGCGCGCGGTTCGCTCAGCTCGT 300
DB 81 ArgProLeuGlySerAlaIleTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100

QY 301 AGACCTACCAAGCTGGGGCGCGCGCTAACCGCGGTTCGCTCGCGGCCCATGACACCCCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120

QY 361 CGAGTGCTGATGCTGACTCGCGCGCGCGCATCTTGGCGCGCGAGTATACCTATCAACA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140

QY 421 TCTCCCTTACCTCTCCGTTGGCCACCGGCACCTAACCTGGTCTTATGCGCGCCCTCTT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160

QY 481 AGTCCGCTTTTACCCCTTCAGGACGGCGCAATACCATATATATGGCCAGCGAGCTTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180

QY 541 AATTATGCCAGATACCGGGTTCGCGCGCGCAATATCGGTATACGCGCGGTGCTGCCAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaAargAlaThrIleArgTyrArgProLeuValProAsn 200

QY 601 GCTGTCGGCGGTACGCGCATCTCCATCTCATTTGGCCACAGACACACACCGCGCGAGC 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220

QY 661 TCCGTTGATATGAATTCATTAACCTCAGCGGATGTCGTATTTTATGTCAGCCCGGCATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240

QY 721 GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATGTATACCAAGGCTGCGCGTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgIleHisTyrArgAsnGlnGlyTrpArgSer 260

QY 781 GTCGAGACCTCTGGGGTGGGTGAGAGAGGCTACCTCTGCTGTGTTATGCTTTTGATA 840
DB 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280

QY 841 CATGGCTCACTCGTAAATTCCTACTATAATACACCTATACCGGTGCCCTCGGGCTGTG 900
DB 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300

QY 901 GACTTTGCCTTGAGTTGAGTTTGCACACCTTACCCCGGTAAACACCAATACCGGGTCC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320

QY 961 TCCCGTTATTCAGCACTGCTCGCACCGCTTGTGCGGTGCGGAGCGGACTGCGGAG 1020
DB 321 SerArgTyrSerSerThrAlaAargHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340

QY 1021 CTCACCAACCGCGTGTACCGGTCTATATGAGGACCTCTATTTTACTAGTACTAATGGT 1080

DB 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360

QY 1081 GTCCGTGAGATCGCGCGGGATAGCCCTCACCTGTTCACCTGTGCTGACACTCTGCTT 1140
DB 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380

QY 1141 GCGCGCTCGCGACAGAAATTTGATTTCGTGCGCTGGTGGCGAGCTGTCTACTCCCGTCCC 1200
DB 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400

QY 1201 GTTGTCTCAGCAATGCGGAGCCGACTGTTAAGTTGTATATCATCTGTAGAGAAATGCTCAG 1260
DB 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420

QY 1261 CAGGATTAAGGTATTSCAATCCCGCATGACATTCACCTCGGAGAAATCTCGTGTGTTATT 1320
DB 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle 440

QY 1321 CAGGATTATGATAACCAACATGAACAGATCGGCGGACGCTTCTCCAGGCCCATCGCGC 1380
DB 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460

QY 1381 CCTTCTCTGCTCTCAGCTAATGATGCTGCTTGGCTCTCTCTCACCCTGCGCGAGTAT 1440
DB 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480

QY 1441 GACCACTCCACTTATGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTCACCTTG 1500
DB 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSer***SerValThrLeu 500

QY 1501 GTTAATCTGTCGCGCGCGCGCGCTTGGCGCGCTCGCTCGATTCGACCAAGGTCAACA 1560
DB 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaAargSerLeuAspTrpThrLysValThr 520

QY 1561 CTTGACGGTCCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTGCTCGCGC 1620
DB 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540

QY 1621 CTCGCGGTGAAGCTCTCTTCTGGAGCGGAGGACCAACTAAAGCGGGTACCTTATAAT 1680
DB 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560

QY 1681 TATAACACCACTGCTAGCGACCACTGCTTGTTCGAGAATCGCGCGCGCACCGGCTGCT 1740
DB 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580

QY 1741 ATTTCCACTTACCACTAGCCTGGGTGCTGGTCCCTCTCCATTTCTGCGGTGCGGTT 1800
DB 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600

QY 1801 TTAGCCCCCACCTCGCTAGCAATTCCTTGAGGATACCTTGGACTACCTGCGCGGCC 1860
DB 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620

QY 1861 CATACTTTTCATGATTTCTGCCAGAGTCCGCCCTTGGCCCTTCAGGCTGCGCTTTC 1920
DB 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640

QY 1921 CAGTCTACTCTGCTGAGCTTCAGCGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 1980
DB 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 13

AA96091
ID AA96091 standard; protein; 549 AA.
XX
AC AA96091;
XX
DT 16-OCT-2003 (revised)
DT 06-AUG-1996 (first entry)
XX
DE Hepatitis E virus (Burma strain) 62K antigen.
XX

KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
KW capsid.
XX
OS Hepatitis E virus; Burma strain.
XX
PN WO9612807-A2.
XX
PD 02-MAY-1996.
XX
PF 23-OCT-1995; 95WO-US013703.
XX
PR 24-OCT-1994; 94US-00327952.
XX
PR 13-OCT-1995; 95US-00542634.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Fuerst TR, Mcatee CF, Yarbough PO, Zhang Y;
XX
DR N-PSDB; AAT27109.
XX
XX
PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
PT diagnostic reagents for determining HEV infection and in vaccines.
XX
PS Claim 3; Page 83-84; 125pp; English.
XX
CC 62K antigen (AAR96091) comprises the C-terminal 549 amino acids of
CC hepatitis E virus (HEV) Burma strain capsid protein (see also AAR96089).
CC It is obt'd. by PCR amplification of HEV Burma ORF-2 (AAT27107), insertion
CC of amplified DNA (AAT27109) into vector pGEX and expression in E. coli
CC cells. Expression of full-length ORF-2 in Sf9 insect cells using a
CC baculovirus system also results in prodn. of 62K antigen (see also
CC AAR96101 and AAR96103). 62K represents an improved antigen, in comparison
CC to bacterial expressed proteins, for use in HEV diagnostic assays, and
CC also has excellent immunogenic properties. (Updated on 16-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 549 AA;

Alignment Scores:
Pred. No.: 0 Length: 549
Score: 549.00 Matches: 549
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.18% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAR96091 (1-549)

QY 334 GCGGTGCGTCCGGCCCATGACACCCGCGAGTGCCTGATGCTCCGGGGCCCATC 393
DB 1 AlaValAlaProAlaHisAspThrProProValProAspValAspSerArgGlyAlaIle 20

QY 394 TTGCGCGCGAGTATAACCTATCAATCTCCCTTACCTCTCCGTGGCCACCGGCAT 453
DB 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaIleThrGlyThr 40

QY 454 AACCTGGTCTTTATGCGCGCCCTTTAGTCGGCTTTTACCCCTTCAGACGGACCAAT 513
DB 41 AsnLeuValLeuTyrAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsn 60

QY 514 ACCCATATAATGGCCACGGAAGCTCTTAATTATGCCAGTACCGGGTTGCCGTCACCA 573
DB 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThr 80

QY 574 ATCCGTTACCGCCGCTGGTCCCAATGCTGTCGGCGGTACGCCATCTCCATCTCATTC 633
DB 81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100

QY 634 TGGCCACAGACACACACCGGCGGCGCTGGATGCAATTCATTAACCTCGACGGAT 693
DB 101 TrpProGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120

QY 694 GTTCGATTTTGTAGTCCAGCCCGCATAGCCTCTGAGCTTGTCATCCCAAGTGAGCGCTTA 753
DB 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140

QY 754 CACTATCGTAAACCAAGGCTGCGCTCCGTCCAGAGCCTCTGGGGTGGCTGAGGAGAGCT 813
DB 141 HisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGluGluAla 160

QY 814 ACCTCTGCTCTTTATGCTTTGTCATACATGGCTCACTCGTAAATTCCTACTACTACA 873
DB 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerLeuValAsnSerTyrThrAsnThr 180

QY 874 CCCTATACCGGTGCGCTCCGTGCTGTGACTTTGCTTCCCTTGAGCTTGAGTTTCGCAACCTT 933
DB 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200

QY 934 ACCCCCGGTAAACCAATACCGGCTCTCCGTTATTCAGACACTGCTGCGCACCGCTT 993
DB 201 ThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220

QY 994 CGTCCGCTGCGGAGCGGACTGCCGAGCTCACCCACGCGGTCTACCGCTTTATGAAG 1053
DB 221 ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLys 240

QY 1054 GACCTCTATTTTACTAGTACTAATGCTGCTGAGATCGGCGCGGAGTAGCCCTCAC 1113
DB 241 AspLeuTyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThr 260

QY 1114 CTGTTCAACCTGCTGACACTCTGCTTGGCGGCTGCCGACAGAAATTAATTTGTCGGCT 1173
DB 261 LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAla 280

QY 1174 GGTGCGCAGCTGTTCTACTCCGCGCTTCTCTCAGCAATGGCGGCGGCTTAAAG 1233
DB 281 GlyGlyGlnLeuPheTyrSerArgProValSerAlaAsnGlyGluProThrValLys 300

QY 1234 TTGTATACATCTGTAGAGAATGCTCAGCAGATAGGGTATTGCAATCCCGCATCACAT 1293
DB 301 LeuTyrThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspIle 320

QY 1294 GACCTCGGAGAAATCTCGTGTGTTATTCAGATTAATGATAACCAACATGAAACAGATCG 1353
DB 321 AspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArg 340

QY 1354 CCGACGCTTCTCCAGCCCATCGCGCTTCTCTGCTCTTCGAGCTAATGATGCTT 1413
DB 341 ProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeu 360

QY 1414 TGGCTCTCTCACCGCTGCCGAGTATGACAGTCCACTTATGGCTCTTCGACTGGCCCA 1473
DB 361 TrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPro 380

QY 1474 GTTTATGTTCTGACTCTGTGACCTTGTGTTAATGTTGGACCGCGCGCGAGCCGCTGCC 1533
DB 381 ValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla 400

QY 1534 CGGTGCTCGATTGGACCAAGTCACTTCAGCGTCCGCCCTCTCCACCATCCAGCAG 1593
DB 401 ArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln 420

QY 1594 TACTCGAAGACCTCTTTGCTGCTCGCGCTCCGCGTAAAGCTCTCTTTCTGGGAGCAGGC 1653
DB 421 TyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGly 440

QY 1654 ACAACTAAGCGGGTACCTTTAATTAATTAACACCACTGCTAGGACCACTGCTTGTTC 1713
DB 441 ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuVal 460

QY 1714 GAGATGCGCGCGGCGGCTGCTATTTCCACTTACACCATGCTAGCTGGTGTGTGT 1773
DB 461 GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGly 480

QY 1774 CCCGTCCTCCATTTCTGGGTTGGCGTTTGTAGCCCCCAGCTCTGGCTAGCATTCCTTGAG 1833
DB 481 ProValSerIleSerAlaValAlaValValAlaProHisSerAlaLeuAlaLeuLeuGlu 500
QY 1834 GATACCTTGAGTACTGCTGCGCGCCGACATATTTTGTATGATTTCTGCCAGAGTGCCTG 1893
DB 501 AspThrLeuAspTyrProAlaArgAlaHisThrPheAspAspPheCysProGluCysArg 520
QY 1894 CCCTTGCGCTTCAGGCTGGCTTTCAGTCTACTCTGCTGAGCTTCAGCGCTTAAG 1953
DB 521 ProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLys 540
QY 1954 ATGAAGCTGGGTAAACTCGGGAGTTG 1980
DB 541 MetLysValGlyLysThrArgGluLeu 549
RESULT 14
AAE12980
ID AAE12980 standard; protein; 660 AA.
XX AC AAE12980;
XX AC AAE12980;
DT 28-JAN-2002 (first entry)
XX Hepatitis E virus ORF2 encoded long capsid protein.
XX Hepatitis E virus; HEV; capsid protein; antiviral; vaccine; therapy;
KW detection; PORF2.
XX Hepatitis E virus.
XX Key Location/Qualifiers
FH Region 221..660
FT /note= "This region contains P1-P51 overlapping synthetic
FT peptides"
XX WO200177156-A2.
XX 18-OCT-2001.
XX 03-APR-2001; 2001WO-US010696.
XX 07-APR-2000; 2000US-0195380P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Fields HA, Khudyakov YE, Meng J;
XX WPI; 2001-657163/75.
XX C-terminal portion of the PORF2 protein of hepatitis E virus polypeptide
PT containing a neutralizing epitope, useful to detect, prevent and treat
PT hepatitis E virus infections.
XX Example; Page 58-61; 62pp; English.
XX The invention relates to a neutralising hepatitis E virus (HEV)
CC polypeptide comprising C-terminal of HEV PORF2 protein containing at
CC least one neutralising epitope. HEV is a non-enveloped virus. HEV
CC polypeptide is useful for studying the pathogenesis of HEV and monitoring
CC treatment efficacy in subjects undergoing treatment for HEV. The
CC polypeptide is used as a reagent to detect HEV and as a vaccine to treat
CC or prevent HEV infection. The present sequence is HEV long capsid protein
CC encoded by open reading frame 2, also referred as PORF2
XX
SQ Sequence 660 AA;
Alignment Scores:
Pred. No.: 0 Length: 660
Score: 541.00 Matches: 641
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 81.97% Indels: 0

DB: 4 Gaps: 0
US-09-851-410a-6_COPY_5147_7129 (1-1983) x AAE12980 (1-660)
QY 55 GCGCCACCGCCCGGTTCAGCGCTTGGCCCGCTGTGGCGGCGGCGAGCGGCGTTCCGGC 114
DB 19 AlaProProGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlySerGly 38
QY 115 GGTGGTTCTTGGGGTGACCGGGTGTATTCTCAGCCCTTCCCAATCCCTATATTCAATCA 174
DB 39 GlyGlyPheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisPro 58
QY 175 ACCAACCCCTTCGCCCGCCGATGTCAACCGCTGCGGCGCGGGCTGACCTCGTTTCGCAA 234
DB 59 ThrAsnProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGln 78
QY 235 CCGCCCGCAGCACTCGGCTCGCTTGGCTGTGACCGCCCGAGCCCGCGGTTGCTCA 294
DB 79 ProAlaArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSer 98
QY 295 CGTGTAGACTACCAACAGCTGGGGCGCGCGCTAACCGGGTCTGCGCGCCCATGAC 354
DB 99 ArgArgArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAsp 118
QY 355 ACCCGCGCAGTGTGATGTGACTCCCGCGCGGCCCATCTTGGCGCGCGCAGTATAACTA 414
DB 119 ThrProValProAspValAspSerArgGlyAlaAlaLeuArgArgGlnTyrAsnLeu 138
QY 415 TCAACATCTCCCTTACCTTCTCGTGGCGCACCGGCACTAACCTGTCTTTATGCGGCC 474
DB 139 SerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAla 158
QY 475 CCTCTAGTCCGCTTTTACCCCTTCAGGAGCGGACCAATACCCATATAATGCGCCGAA 534
DB 159 ProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGlu 178
QY 535 GCTTCTAATATATGCCAGTACCGGGTTGCGCGTGCACAAATCCGTTTACCGCCGCTGTC 594
DB 179 AlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuVal 198
QY 595 CCAATGCTGTCGGCGGTATACGCATCTCATCTTCTTGGCCACAGACACACACACC 654
DB 199 ProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThr 218
QY 655 CCGACGTCCGTTGATATGAATTCATTAACCTCGACGATGCTCGTATTTAGTCCAGCC 714
DB 219 ProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnPro 238
QY 715 GGCATAGCTCTGAGCTTGTGATCCCAAGTGAGCGCCTACACTATCTGTAACCAAGGTGG 774
DB 239 GlyIleAlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyr 258
QY 775 CGCTCCGTCGAGACCTCTGGGGTGGCTGAGGAGAGGCTACCTCTGCTTGTATTGCTT 834
DB 259 ArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeu 278
QY 835 TGCATACATGGCTCACTCGTAATTCCTATATACACCTATACCTATACCGTCCCTGGG 894
DB 279 CysIleHisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGly 298
QY 895 CTGTGGACCTTTCGCCCTTGAGCTTGTGATTCGCAACCTTACCCCGGTAAACCAATACG 954
DB 299 LeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThr 318
QY 955 CCGGTCTCCCGTTATTCCAGCACTGTGCGCCACCGCTTCTGTCGGGTGCGGACGGGACT 1014
DB 319 ArgValSerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThr 338
QY 1015 GCGAGCTCACACACCGCTGCTACCGCTTTATGAGACCTCTATTATTACTAGTACT 1074
DB 339 AlaGluLeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThr 358
QY 1075 AATGGTGTGCTGAGTATCGCGGGGATAGCCCTCACCCTGTTTCAACCTTGTGCTGACACT 1134

```
|||||
Db 359 AsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThr 378
QY 1135 CTGCTGGCGGCTCCGACAGAAATGATTTCGTGGGTGGGCGAGCTGTTCTACTCC 1194
Db 379 LeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyCysGlnLeuPheTyrSer 398
QY 1195 GTCGCGCTGCTCAGCAATGGCGAGCGACTGTTAAGTTGATATACATCTCTAGAGAT 1254
Db 399 ArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsn 418
QY 1255 GCTCAGCAGGATTAAGGATTGCAATCCGCGATGATTCGAGCTCGGAGAATCTCGTGTG 1314
Db 419 AlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgVal 438
QY 1315 GTTATTCAGGATTATGATAACCAACATGAACAGATCGGCCGACGCTTCTCCAGCCCA 1374
Db 439 ValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaPro 458
QY 1375 TCGCGCCCTTCTCTGCTCTCGAGCTAATGATGTGCTTTGGCTCTCTCACCGCTGCC 1434
Db 459 SerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAla 478
QY 1435 GAGTATGACCACTCCACTATGGCTCTTCGACTGGCCAGTTATGTTTCTGACTCTGTG 1494
Db 479 GluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerVal 498
QY 1495 ACCTTGGTTAATGTTCGACCGCGCGAGCGGCTGCTCCGCTCGCTCGATGGACCAAG 1554
Db 499 ThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLys 518
QY 1555 GTCACACTTGACGTCGCGCCCTCTCCACCATCCAGAGTACTCGAAGACCTTCTTTGTC 1614
Db 519 ValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheVal 538
QY 1615 CTGCGCTCCGCGGTAAGCTCTTTCTGGGAGGCGAGCACAACTAAAGCCGGTACCTC 1674
Db 539 LeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrPro 558
QY 1675 TATATTATACACCACTGCTAGCCAGCACTGCTGTGTCGAAATGCCCGCGGACCGG 1734
Db 559 TyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArg 578
QY 1735 GTCGCTATTTCACCTACACCACTAGCTGGGTGCTGCTCCGCTCTCCATTTCTCGGTT 1794
Db 579 ValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSerAlaVal 598
QY 1795 GCGGTTTTCAGCCCACTCTGCGTAGCATTCGTGAGGATACCTTGAGTACCTGCGC 1854
Db 599 AlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAla 618
QY 1855 CGCGCCCATCTTTTGATGATTCTGCGCAGAGTCCGCGCCCTTGGCTTCAGGCTGC 1914
Db 619 ArgAlaHisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCys 638
QY 1915 GCTTTCAGTCTVACTGTCGCTGAGCTTCAGCGCCCTTAAGATGAAGGTGGGTAAACTCGG 1974
Db 639 AlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArg 658
QY 1975 GAGTTG 1980
Db 659 GluLeu 660
```

RESULT 15

AAR96101

ID AAR96101 standard; protein; 540 AA.

XX AC

AAR96101;

DT 16-OCT-2003 (revised)

DT 06-AUG-1996 (first entry)

XX XX

DE Hepatitis E virus (Burma strain) recombinant 62K antigen.

```
XX
KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
KW capsid.
XX
OS Hepatitis E virus; Burma strain.
XX
PN WO9612807-A2.
XX
PD 02-MAY-1996.
XX
PF 23-OCT-1995; 95WO-US013703.
XX
PR 24-OCT-1994; 94US-00327952.
PR 13-OCT-1995; 95US-00542634.
XX
(PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;
XX
WPI; 1996-230608/23.
XX
Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
diagnostic reagents for determining HEV infection and in vaccines.
XX
Claim 5; Page 92-93; 125pp; English.
XX
DNA (AAT27109) coding for the 62K antigen (AAR96091) of hepatitis E virus
(HEV) Burma strain capsid protein was cloned into baculovirus expression
vector pBluscriptII and recombinant 62K was expressed in Sf9 insect cells.
High levels of expression were obtained, and the recombinant 62K was obtained
over 95% purity. However, C-terminal processing resulted in the deletion
of 9 or 23 amino acids from 62K, giving 2 related polypeptide species
(AAR96101 and AAR96103). Similar results were obtained with HEV Mexico 62K
antigen (see also AAR96102 and AAR96104). Recombinant 62K represents an
improved antigen, in comparison to bacterial expressed proteins, for use
in HEV diagnostic assays, and also has excellent immunogenic properties.
(Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 540 AA;
```

Alignment Scores:

```

Pred. No.: 0 Length: 540
Score: 540.00 Matches: 540
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 81.82% Indels: 0
DB: 2 Gaps: 0
```

US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAR96101 (1-540)

```

QY 334 GGGTTCGCTCCGCGCCCATGACACCCGCGAGTGCCTGATGCTGACTCCGCGGCGCCATC 393
Db 1 AlaValAlaProAlaHisAspThrProValProAspValAspSerArgGlyAlaIle 20
QY 394 TTGCGCCGCGCAGTATACACTCTCCCTTACCTCTCCGTCGCCACCGGCACT 453
Db 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThr 40
QY 454 AACCTGGTTCCTTATGCGCGCCCTCTTAGTCGCTTTTACCCTTCAGGACGGACCAAT 513
Db 41 AsnLeuValLeuTyrAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACCCATATAATGGCCACGGAAGCTCTTAATATGCCAGTACCGCGTTGCCGTCGCCACA 573
Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaAlaThr 80
QY 574 ATCCGTTTACCGCGCGCTGTCGCCAATGCTGTGCGCGGTAGCCCATCTCCATCTCATTC 633
Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100
QY 634 TGGCCACAGACCCACCACCCGCGCTCCGTTGATATGAATTCATACCTTCGCGGAT 693
|||||
```

```
Db 101 TrpProGlnThrThrThrThrThrThrSerValAspMetAsnSerIleThrSerThrAsp 120
QY 694 GTTGGTATTTAGTCCAGCCCGGATAGCCTCTGAGCTTGATCCCAAGTGAAGCGCTA 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTATCGTAACCAAGGTGGCGTCCCGTCCAGACCTCTGGGGTGGCTGAGGAGGAGGCT 813
Db 141 HisTyrArgAsnGlnGlyTyrPargSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCTGTCTGTGTATGCTTTGTCATACATCGCTCAGTCTGCTAAATTCCTATCTACTAATACA 873
Db 161 ThrSerGlyLeuValMetCysIleHisGlySerLeuValAsnSerTyrThrAsnThr 180
QY 874 CCCTATACCGGTGCCCTGGGCTGTGGACTTTCGCCCTTGAGCTTGAGCTTTCGCAACCTT 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuPheArgAsnLeu 200
QY 934 ACCCGCGGTAAACCAATACCGGGTCTCCCGTTATTCAGCACTGCTCGCCACCGCCTT 993
Db 201 ThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
QY 994 CGTCGCGGTGGAGCGGAGTCCGAGCTCACACACGGTGTCTACCCGCTTTATGAAG 1053
Db 221 ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLys 240
QY 1054 GACCTCTATTTTACTAGTACTAATGGTGTGGTGAGATCGCGCGGATAGCCCTCACC 1113
Db 241 AspLeuTyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThr 260
QY 1114 CTGTTCAACCTTGCTGACACTCTGCTTGGCGGCTTCCGCGAGCAATTCGATTCGTCGGCT 1173
Db 261 LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAla 280
QY 1174 GGTGGCCAGCTGTCTACTCCCGTCCCGTGTCTCAGCCCAATGGCGACCGCACTGTTAAG 1233
Db 281 GlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLys 300
QY 1234 TTGTATACATCTGTAGAGAATGCTCAGCAGGATAGGGTATTCGATCCCGCATGACATT 1293
Db 301 LeuTyrThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspIle 320
QY 1294 GACCTCGGAGAACTCTGTGTGTTATTCAGATTATGATACCAACATGAACAAGATCGG 1353
Db 321 AspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArg 340
QY 1354 CCGACGCCCTTCTCCAGCCCGCATCGCGCCTTCTCTGTCTCGAGCTAATGATGTGCTT 1413
Db 341 ProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeu 360
QY 1414 TGGCTCTCTCTCAGCGTCCGAGTATGACAGTCCACTATGCTCTTCGACTGGCCCA 1473
Db 361 TrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerThrGlyPro 380
QY 1474 GTTATGTTCTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGCAGGCCGCTTGCC 1533
Db 381 ValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla 400
QY 1534 CGGTGCTCGATTGACCAAGTCACTTGAAGTCCGCCCTCTCCACCATCCAGCAG 1593
Db 401 ArgSerLeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln 420
QY 1594 TACTCGAGACCTTCTTTGTCTGCGGCTCCGCGTAAAGTCTCTTTCTGGGAGGCAGGC 1653
Db 421 TyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGly 440
QY 1654 ACAACTAAAGCCGGGTACCTTTATATTATATACCACTGCTAGCGACCACTGCTGTC 1713
Db 441 ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuVal 460
QY 1714 GAGAAATGCCCGGCGGCGGCTGCTATTTCCACTTACCACTAGCCTGGGTGCTGCT 1773
Db 461 GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGly 480
```

Search completed: August 16, 2004, 13:47:46
Job time : 126.818 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:40:03 ; Search time 31.2737 Seconds
(without alignments)
12198.625 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129
Perfect score: 660
Sequence: 1 ATGCGCCCTCGGCTATT...GTAAACTCGGAGTTGTAG 1983

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgm21/USPTO_spool_p/US09851410/runat_16082004_125717_9653/app_query.fasta_1.4238
-DB=PIR_78 -QFMT=fastan -SUFFIX=oli.rpr -MINMATCH=0.1 -LOOPEXT=0
-UNITS=BITS -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09851410@CGN_1_1_101@runat_16082004_125717_9653 -NCPU=6 -ICPU=3
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	660	100.0	660	1 VHWWH2	structural protein
2	120	18.2	659	1 B44212	structural protein
3	109	16.5	123	1 VHWHE	structural protein
4	61	9.2	123	1 C44212	structural protein
5	14	2.1	41	1 B40236	structural protein
6	9	1.4	337	2 B86242	hypothetical prote
7	9	1.4	348	2 T12281	NADH2 dehydrogenas
8	9	1.4	840	2 H96635	unknown protein, 7
9	9	1.4	1091	2 F83928	hypothetical prote
10	8	1.2	41	2 H58208	protamine II - bia
11	8	1.2	74	2 S19708	hypothetical prote
12	8	1.2	97	2 D81144	hypothetical prote
13	8	1.2	138	2 JT0593	hypothetical 15K p
14	8	1.2	166	2 D72394	hypothetical prote

15	8	1.2	172	2 F87649	ExdD/TolR family p
16	8	1.2	199	2 S05509	photosystem II oxy
17	8	1.2	202	2 S72224	oxygen-evolving en
18	8	1.2	220	2 T50624	hypothetical prote
19	8	1.2	222	2 H96711	hypothetical prote
20	8	1.2	226	2 C97488	hypothetical 21.6K
21	8	1.2	231	2 D82555	heme ABC transport
22	8	1.2	233	1 WMBEHK	UL3 protein - huma
23	8	1.2	239	2 AD3571	transcriptional regu
24	8	1.2	240	2 T23797	hypothetical prote
25	8	1.2	241	2 S73170	ribosomal protein
26	8	1.2	252	2 T45737	hypothetical prote
27	8	1.2	266	2 H83564	probable short-cha
28	8	1.2	273	2 T48639	hypothetical prote
29	8	1.2	277	2 A87184	thiosulfate sulfur
30	8	1.2	292	2 H70313	cobalamin synthesi
31	8	1.2	309	2 S10889	proline-rich prote
32	8	1.2	318	2 T00402	homeodomain transc
33	8	1.2	319	2 F82601	5'-nucleotidase XF
34	8	1.2	339	2 C83290	probable transcrip
35	8	1.2	344	2 T14018	NADH2 dehydrogenas
36	8	1.2	344	2 T17077	NADH2 dehydrogenas
37	8	1.2	345	2 T12365	NADH2 dehydrogenas
38	8	1.2	345	2 T12366	NADH2 dehydrogenas
39	8	1.2	345	2 T17053	NADH2 dehydrogenas
40	8	1.2	358	2 S58376	aryl hydrocarbon r
41	8	1.2	374	2 T20481	hypothetical prote
42	8	1.2	397	2 S27600	N-acetylglutamate
43	8	1.2	397	2 F86680	F5114.10 (imported
44	8	1.2	412	2 F88563	protein T05G5.1 li
45	8	1.2	442	1 S11712	transcription init

ALIGNMENTS

RESULT 1

VHWWH2

structural protein 2 precursor - hepatitis E virus (strain Burma)

C:Species: hepatitis E virus

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 23-Jul-1999

C:Accession: C40778

R:Tam, A.W.; Smith, M.M.; Guerra, M.E.; Huang, C.C.; Bradley, D.W.; Fry, K.E.; Reyes, G.

Virolology 185, 120-131, 1991

A:Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length vi

A:Reference number: A40778; PMID:92024067; PMID:1926770

A:Accession: C40778

A:Molecule type: Genomic RNA

A:Residues: 1-660 <TAM>

A:Cross-references: GB:M73218; NID:g330023; PIDN:AAA45736.1; PID:g330026

A>Note: the authors translated the codon CGC for residue 2 as Ala

C:Superfamily: hepatitis E virus structural protein 2

C:Keywords: structural protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-660/Product: structural protein 2 #status predicted <SP2>

Alignment Scores:

Pred. No.:	Score:	Length:
0	660.00	660
Percent Similarity:	100.00%	Matches: 0
Best Local Similarity:	100.00%	Conservative: 0
Query Match:	100.00%	Mismatches: 0
DB:	1	Indels: 0
		Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x VHWWH2 (1-660)

QY 1 ATGCGCCCTCGGCTATTTTTGTGCTCTCATGTTTTTTCCTATGTCGCCGCGCA 60

|||||

1 MetAProArgProIleLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20

QY 61 CGCCCGGTACAGCTCTGGCGCGCGCGAGCGGGTTCGCGCGGTGT 120

|||||

21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlyGlyGly 40

121 TTCTGGGTTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 180
Db |||||||
41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY |||||||
181 CCCTTCGCCGCCGATGTACCGCTGCGCGCGGGCTGGACCTCGTGTTCGCCAACCGGCC 240
Db |||||||
61 ProPheAlaProAspValThrAlaAlaAlaGlyProArgValArgGlnProAla 80
QY |||||||
241 CGACCACTCGCTCCGCTCGGTGGTGACAGGCCAGCGCCCGCGGTGGCTCAGTGTGT 300
Db |||||||
81 ArgProLeuGlySerAlaTrpArgGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY |||||||
301 AGACCTACCAACAGCTGGCGCGCGCGCTAACCGCGTGGCTCGCGGCCCATGACACCCG 360
Db |||||||
101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY |||||||
361 CCAGTGCCTGAATCGACTCCCGCGCGCATCTTGGCGCGGAGTATACCTATCAACA 420
Db |||||||
121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY |||||||
421 TCTCCCTTACCTCTTCGTTGGCCACCGGCACTAACTGGTCTCTTATGCGCGCCCTCT 480
Db |||||||
141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY |||||||
481 AGTCCGCTTTTACCCCTTCAGCAGCGGCACCAATACCATATATATGGCCACGGAAGCTTCT 540
Db |||||||
161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY |||||||
541 AATTATGCCAGTACCGGGTGGCCGTCGCACCAATCCGTTACCGCGCGTGGTCCCAAT 600
Db |||||||
181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY |||||||
601 GCTGTGGCGGGTTACGCCATCTCCATCTCATCTTTCGCCACAGACCACCCCGGACG 660
Db |||||||
201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY |||||||
661 TCCGTTGATATGAATCAATAACTCGACGATGTTGCTATTTTAGTCCAGCCCGCATA 720
Db |||||||
221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY |||||||
721 GCCTCTGAGCTTGATGCCAAGTACGCGCTACACTATGTAACCAAGCTGGCGCTCC 780
Db |||||||
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY |||||||
781 GTCCAGACCTCTGGGTGGTGGAGGAGGCTACCTCTGCTTCTGTTATGCTTTCATA 840
Db |||||||
261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY |||||||
841 CATGGCTCACTCGTAATTCCTATACATAACACCTATACCGGTGCGCTCGGGCTGTG 900
Db |||||||
281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY |||||||
901 GACTTTGCCCTTGAGCTTGATTTGCAACCTTACCCCGGTAAACCAATACGCGGTC 960
Db |||||||
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY |||||||
961 TCCCGTTATTCAGCACTGCTCCACCGCTTCGTCGGGTGGCGGACCGGCTGGCGAG 1020
Db |||||||
321 SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY |||||||
1021 CTCACCAACACCGCTCTACCGCTTATGAAGACCTCTATTTTACTAGTACTAATGTT 1080
Db |||||||
341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY |||||||
1081 GTCCGTGAGATCGCCGCGGATAGCCCTCACCTGTTCAACCTTCTGCTGACACTGTGTT 1140
Db |||||||
361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY |||||||
1141 GCGCGCTCGCCACAGATTGTTTCGTGGTGGTGGCCAGCTGTTCTACTCCCGTCCC 1200
Db |||||||
381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlnLeuPheTyrSerArgPro 400
QY |||||||
1201 GTTGTCTCAGCCAATGGCGCGCACTGTTTAAAGTTGTATATACATCTCTAGAGAATGCTCAG 1260

Db |||||||
401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY |||||||
1261 CAGGATAAGGGTATTCAATCCGCGATGACATTGACCTCGAGAATCTCGTGGTTATT 1320
Db |||||||
421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle 440
QY |||||||
1321 CAGGATTATGATAACCAACATGAACAGATCGCGCGCGCGCTTCTCAGCCCCATCGCGC 1380
Db |||||||
441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY |||||||
1381 CCTTCTCTCTCTCGAGCTAATGATGTCTTGGCTCTCTCTCACCCTCGCGAGTAT 1440
Db |||||||
461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY |||||||
1441 GACCACTCCACTTATGGCTTCTCGACTGGCCCCAGTTATGTTTCTGACTCTGACCTTG 1500
Db |||||||
481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY |||||||
1501 GTTAATGTTGCCAGCGCGCGCGCGTGGCCGCTCGCTCGATTGGACCAAGTCA 1560
Db |||||||
501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY |||||||
1561 CTTGACGCTCGCCCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTCGCGC 1620
Db |||||||
521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY |||||||
1621 CTCGCGGTAAGCTCTCTTTCGGGAGCGAGGCACAACTAAAGCCGGGTACCTTTAAT 1680
Db |||||||
541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY |||||||
1681 TATAACCACTCTAGCAGCACTCTTGTTCGAGAAATGCCCGCGGACCCGGTCCGCT 1740
Db |||||||
561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY |||||||
1741 ATTTCCACTTACCACTAGCCTGCTGGTGGTCCGCTCTCCATTTCTCGGTTGCCGTT 1800
Db |||||||
581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY |||||||
1801 TTAGCCCCCACTCTCGCTAGCATTTGTTGAGGATACCTTGGACTACCTCGCCGCGCC 1860
Db |||||||
601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY |||||||
1861 CATACCTTTGATGATTTCTGCCAGAGTCCCGCCCTTGGCTTCCAGGCTCGGCTTTC 1920
Db |||||||
621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY |||||||
1921 CAGTCTACTGCTCGCTCAGCTTACGCGCTTAAAGATGAAGTGGGTAAACCTCGGAGTTG 1980
Db |||||||
641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 2

B44212

structural protein 2 precursor - hepatitis E virus (strain Mexico)

C:Species: hepatitis E virus

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999

C:Accession: B44212

R:Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.

Virology 191, 550-558, 1992

A:Title: Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HE)

A:Reference number: B44212, MUID:93079857, PMID:1448913

A:Accession: B44212

A:Molecule type: genomic RNA

A:Residues: 1-659 <HUA>

A:Cross-references: GB:M74506; NID:9330017; PIDN:AAA45732.1; PID:9330020

C:Superfamily: hepatitis E virus structural protein 2

C:Keywords: structural protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-659/Product: structural protein 2 #status predicted <SP2>

Alignment Scores:

Pred. No.: 2,05e-109

Score: 120.00

Length: 659

Matches: 120

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.18% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x B44212 (1-659)

QY 490 TTACCCCTTCAGCGGACCAATACCCATATAATGGCCACGAAAGCTTCTAATTATGCC 549
DB 164 LeuProLeuGlnAspGlyThrAsnThrHisLeuMetAlaThrGluAlaSerAsnTyrAla 183
QY 550 CAGTACCGGGTTCGCCGCTCCCAATTCGTTACCCCGCTGGTCCCAATGCTCTCGGC 609
DB 184 GlnTyrArgValAlaAraGAlaThrileArgTyrArgProLeuValProAsnAlaValGly 203
QY 610 GGTTACGCCATCTCCATCTCTTGGCCACAGACACACCCCGACGTCGGTTGAT 669
DB 204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCATAACCTCCAGCGATGTCGTATTTAGTCCAGCCCGCATAGCCTCTGAG 729
DB 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
QY 730 CTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGCTCCGTCGAGACC 789
DB 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThr 263
QY 790 TCTGGGGTGGCTGAGGAGGAGGCTACCTCTGTCTTGTATGCTTTGCATACATGGCTCA 849
DB 264 SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283

RESULT 3

VHWHE

structural protein 1 - hepatitis E virus (strain Burma)

C:Species: hepatitis E virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 23-Jul-1999
C:Accession: B40778; A40236
R:Tam, A.W.; Smith, M.M.; Guerra, M.E.; Huang, C.C.; Bradley, D.W.; Fry, K.E.; Reyes, G.
Virology 185, 120-131, 1991
A:Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length vi
A:Reference number: A40778; PMID:92024067; PMID:1926770
A:Accession: B40778
A:Molecule type: genomic RNA
A:Residues: 1-123 <TAM>
A:Cross-references: GB: M73218; NID: g330023; PIDN: AAA45735.1; PID: g330025
R:Ray, R.; Jameel, S.; Manivel, V.; Ray, R.
Virology 189, 359-362, 1992
A:Title: Indian hepatitis E virus shows a major deletion in the small open reading frame
A:Reference number: A40236; PMID:92295577; PMID:1534953
A:Accession: A40236
A:Molecule type: genomic RNA
A:Residues: 1-57, 'p', 59-102, 'p', 104-123 <RAY>
C:Superfamily: hepatitis E virus structural protein 1
C:Keywords: structural protein

Alignment Scores:
Pred. No.: 1 85e-98 Length: 123
Score: 109.00 Matches: 109
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.52% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x VHWHE (1-123)

QY 2 TGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTTGCTATGCTGCGCGGCCAC 61
DB 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGCCGGTTCAGCGCTGCGCGCTGCGGCGGCGGCGGCGGCTTCGCGGGTGGTT 121
DB 35 ArgProValSerArgLeuAlaAlaValValGlyGlyAlaAlaValProAlaValVal 54

QY 122 TCTGGGGTACCGGGTTCATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAACC 181
DB 55 SerGlyValThrGlyLeuIleLeuSerProSerGlnSerProIlePheIleGlnProThr 74
QY 182 CTTTCGCCCGCATGTCACCGCTGCGCGCGGGCTGGACCTCGTGTTCGCCAACCGGCC 241
DB 75 ProSerProProMetSerProLeuArgProGlyLeuAspLeuValPheAlaAsnProPro 94
QY 242 GACCACTCGGCTCCGCTTGGGTGACAGGCCAGGCCCGCCCGCTGCTCCTCAGCTCGTA 301
DB 95 AspHisSerAlaProLeuGlyValThrArgProSerAlaProProLeuProHisValVal 114
QY 302 GACCTACCAACAGCTGGGGCGCGCGCGC 328
DB 115 AspLeuProGlnLeuGlyProArgArg 123

RESULT 4

C44212

structural protein 1 - hepatitis E virus (strain Mexico)

C:Species: hepatitis E virus

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999

C:Accession: C44212

R:Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.
Virology 191, 550-558, 1992

A:Title: Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HE

A:Reference number: A44212; PMID:93079857; PMID:1448913

A:Accession: C44212

A:Molecule type: genomic RNA

A:Residues: 1-123 <HUA>

A:Cross-references: GB: M74506; NID: g330017; PIDN: AAA45731.1; PID: g330019

C:Superfamily: hepatitis E virus structural protein 1

C:Keywords: structural protein

Alignment Scores:
Pred. No.: 4 45e-51 Length: 123
Score: 61.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.24% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x C44212 (1-123)

QY 2 TGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTTGCTATGCTGCGCGGCCAC 61
DB 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGCCGGTTCAGCGCTGCGCGCTGCGGCGGCGGCGGCGGCTTCGCGGGTGGTT 121
DB 35 ArgProValSerArgLeuAlaAlaValValGlyGlyAlaAlaValProAlaValVal 54
QY 122 TCTGGGGTACCGGGTTCATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAACC 181
DB 55 SerGlyValThrGlyLeuIleLeuSerProSerGlnSerProIlePheIleGlnProThr 74

QY 182 CCT 184

DB 75 Pro 75

RESULT 5

B40236

structural protein 1 - hepatitis E virus (strain Indian)

C:Species: hepatitis E virus

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999

C:Accession: B40236

R:Ray, R.; Jameel, S.; Manivel, V.; Ray, R.

Virology 189, 359-362, 1992

A:Title: Indian hepatitis E virus shows a major deletion in the small open reading frame

A:Reference number: A40236; PMID:92295577; PMID:1534953

A:Accession: B40236

A:Molecule type: genomic RNA

A:Residues: 1-41 <RAY>

C:Superfamily: hepatitis E virus structural protein 1

C;Keywords: structural protein

Alignment Scores:
Pred. No.: 0.000128 Length: 41
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.12% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x B40236 (1-41)

QY 287 TTGCTCACGCTAGACCTACACAGCTGGGCGCGCGC 328

|||||

Db 28 LeuProHisValIaspIeuProGlnIeuGlyProArgArg 41

RESULT 6

B86242

hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: B86242

R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B86242

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-337 <STO>

A;Cross-references: GB:AE005172; NID:g4874264; PIDN:AAD31329.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

Alignment Scores:
Pred. No.: 8.28 Length: 337
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.39% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x B86242 (1-337)

QY 1953 CTTAAGCGCTGAAGCTCAGCGACGT 1927

|||||

Db 149 LeuIysAlaLeuIysIeuSerAspSer 157

RESULT 7

T12281

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - Poecilia sphenops mitochondrion

C;Species: mitochondrion Poecilia sphenops

C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 03-Jun-2002

C;Accession: T12281

R;Breden, F.; Pracek, M.; Rashed, M.; Taphorn, D.; de Figueiredo, C.A.

submitted to the EMBL Data Library, October 1997

A;Description: Molecular phylogeny of a live-bearing fish genus Poecilia (Cyprinodontifc

A;Reference number: Z17479

A;Accession: T12281

A;Status: preliminary; translated from GB/EMBL/DDBU

A;Molecule type: DNA

A;Residues: 1-348 <BRE>

A;Cross-references: EMBL:AF031390; NID:g2653584; PID:g2653585; PIDN:AAB87615.1

C;Genetics:

A;Genome: mitochondrion

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 2

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Alignment Scores:
Pred. No.: 8.24 Length: 348
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x T12281 (1-348)

QY 1910 GCTGCGCTTTCAGTCTACTGTGCTG 1936

|||||

Db 281 AlaAlaLeuSerIeuLeuSerIeu 289

RESULT 8

H96635

unknown protein, 73054-77165 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: H96635

R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: H96635

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-840 <STO>

A;Cross-references: GB:AE005173; NID:g6751691; PIDN:AAF27674.1; GSPDB:GN00141

C;Genetics:

A;Gene: T7P1.16

A;Map position: 1

Alignment Scores:
Pred. No.: 7.31 Length: 840
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x H96635 (1-840)

QY 1907 AGGCTGCGCTTTCAGTCTACTGTGCG 1933

|||||

Db 114 ArgAlaAlaLeuSerIeuLeuSer 122

RESULT 9

F83928

hypothetical protein BH2230 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C;Accession: F83928

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: F83928

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1091 <STO>

A;Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BAB05949.1; GSPDB:GN00

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH2230

Alignment Scores:
Pred. No.: 7.06 Length: 1091
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x F83928 (1-1091)

QY 19 TTGTTGCTGCTCTCATGTTTTCCT 45

Db 12 LeuLeuLeuLeuLeuMetPheLeuPro 20

RESULT 10

H58208

protamine II - black rat snake (fragment)

C;Species: Elaphe obsoleta quadrivittata

C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 07-May-1999

C;Accession: H58208

R;Hunt, J.G.; Kasinsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp, D.

J. Biol. Chem. 271, 23547-23557, 1996

A;Title: Protamines of reptiles.

A;Reference number: A58208; MUID:96394458; PMID:8798564

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-41 <HUN>

C;Superfamily: sperm histone

Alignment Scores:

Pred. No.: 107 Length: 41
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x H58208 (1-41)

QY 76 TCTGGCCGCGTGTGGGCGGCG 99

Db 13 SerGlyArgArgArgGlyArg 20

RESULT 11

S19708

hypothetical protein C - Herpetosiphon aurantiacus

C;Species: Herpetosiphon aurantiacus

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Jan-1995

C;Accession: S19708

R;Erdmann, D.; Duesterhoeft, A.; Kroeger, M.

Eur. J. Biochem. 202, 1247-1256, 1991

A;Title: Cloning and molecular characterization of the HgiCI restriction/ modification

A;Reference number: S19706; MUID:92111503; PMID:1662609

A;Accession: S19708

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-74 <KRO>

A;Cross-references: EMBL:X55138

Alignment Scores:

Pred. No.: 98.7 Length: 74
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x S19708 (1-74)

QY 827 ACAAGACCAGAGGTAGCCTCCTCC 804

Db 64 ThrArgProGluValAlaSerSer 71

RESULT 12

D81144

hypothetical protein NMB0903 [imported] - Neisseria meningitidis (strain MC58 serogroup C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C;Accession: D81144

R;Tetzelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: A81000; MUID:20175755; PMID:10710307

A;Accession: D81144

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-97 <TET>

A;Cross-references: GB:AE002442; GB:AE002098; NID:g7226133; PIDN:AAF41311.1; PID:g722614

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB0903

Alignment Scores:

Pred. No.: 95.2 Length: 97
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x D81144 (1-97)

QY 837 GCAAAGCATACAAACACAGAGGT 814

Db 43 AlaLysHisAsnLysThrArgGly 50

RESULT 13

JT0593

hypothetical 15K protein (hgidiim 3' region) - Herpetosiphon aurantiacus

N;Alternate names: ORP15

C;Species: Herpetosiphon aurantiacus

C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 04-Mar-2000

C;Accession: JT0593; S21948

R;Duesterhoeft, A.; Kroeger, M.

Gene 106, 87-92, 1991

A;Title: Cloning, sequence and characterization of m5C-methyltransferase-encoding gene,

A;Reference number: JT0592; MUID:92039068; PMID:1937045

A;Accession: JT0593

A;Molecule type: DNA

A;Residues: 1-138 <DUE>

A;Cross-references: EMBL:X55141; NID:g48771; PIDN:CAA38940.1; PID:g48772

A;Experimental source: strain Hpa2

C;Superfamily: Herpetosiphon aurantiacus hypothetical 15K protein (hgidiim 3' region)

Alignment Scores:

Pred. No.: 90.7 Length: 138
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x JT0593 (1-138)

QY 822 ACCAGAGTAGCCTCCTCCTCAGC 799

Db 43 ThrArgGlySerLeuLeuSer 50

RESULT 14

D72394

hypothetical protein TM0292 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: D72394
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: D72394
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-166 <ARN>
A;Cross-references: GB:AE001711; GB:AE000512; NID:g4980788; PIDN:AAD35380.1; PID:g498079
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0292

Alignment Scores:
Pred. No.: 88.5 Length: 166
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x D72394 (1-166)

Qy 1728 GCACGGGTGCTATTTCCACTTA 1751

Db 19 AlaProGlyArgTyrPheHisLeu 26

RESULT 15

F87649
ExbD/TolR family protein [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: F87649
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: F87649
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-172 <STO>
A;Cross-references: GB:AB005673; NID:g13424916; PIDN:AAK25194.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC3232

Alignment Scores:
Pred. No.: 88 Length: 172
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x F87649 (1-172)

Qy 79 GGCGCGGTGCTGGCGCGCAGC 102

Db 16 GlyArgArgGlyArgGser 23

Search completed: August 16, 2004, 13:57:39
Job time : 40.2737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: August 16, 2004, 13:36:13 ; Search time 18.2638 Seconds
(without alignments)
11307.070 Million cell updates/sec

Title: US-09-851-410a-6_COPY_5147_7129
Perfect score: 660
Sequence: 1 ATGGCGCTCGGCTATTTT.....GTAAACTCGGAGTTGTAG 1983

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues
Word size: 1
Total number of hits satisfying chosen parameters: 283186

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO_spool_p/US09851410/runat_16082004_125716_9625/app_query.fasta_1.4238
-DB=SwissProt 42 -QFMT=fastan -SUFFIX=oli.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09851410@cgn 1.1.43 @runat_16082004_125716_9625 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	660	100.0	660	1 VST2 HEVBU	P29326 hepatitis e
2	408	61.8	660	1 VST2 HEVPA	P33426 hepatitis e
3	128	19.4	660	1 VST2 HEVHY	Q04611 hepatitis e
4	120	18.2	485	1 VST2 HEVRH	Q00270 hepatitis e
5	120	18.2	659	1 VST2 HEVME	Q03500 hepatitis e
6	109	16.5	123	1 VST1 HEVBU	P29325 hepatitis e
7	76	11.5	123	1 VST1 HEVHY	Q04612 hepatitis e
8	61	9.2	123	1 VST1 HEVME	Q03499 hepatitis e
9	9	1.4	337	1 IF33 ARATH	Q9C522 arabidopsis
10	8	1.2	121	1 RL7 CLOPE	Q8xhr7 clostridium
11	8	1.2	138	1 YD2M HERAU	P25279 herpetosiph
12	8	1.2	166	1 LED1 THEMA	Q9wyc8 thermotoga
13	8	1.2	199	1 PSBQ CHLRE	P12852 chlamydomon
14	8	1.2	202	1 PSBQ VOLCA	Q41643 volvox cart
15	8	1.2	218	1 LOLE XANCP	Q8pc65 xanthomonas
16	8	1.2	233	1 UL03 HSV2H	P28279 herpes simp
17	8	1.2	241	1 RR2 FORPU	P51249 porphyra pu
18	8	1.2	258	1 VSP3 TRIGA	Q13063 trimeresuru

19	8	1.2	273	1 DEFC ARATH	Q9fuz2 arabidopsis
20	8	1.2	277	1 THTR MYCLE	Q50036 mycobacteri
21	8	1.2	318	1 ATH4 ARATH	P92953 arabidopsis
22	8	1.2	331	1 MAN1 MOUSE	Q9wu40 mus musculu
23	8	1.2	358	1 PTPA HUMAN	Q15257 homo sapien
24	8	1.2	412	1 YNP1 CAEEL	P34554 caenorhabdi
25	8	1.2	507	1 MKR3 HUMAN	Q13064 homo sapien
26	8	1.2	511	1 HRDB STRCO	P18183 streptomyce
27	8	1.2	540	1 SNB2 HUMAN	Q13425 homo sapien
28	8	1.2	634	1 ICPG SYNY3	P37979 synechocyst
29	8	1.2	732	1 PSAB CYAME	Q85fy6 cyanidiosch
30	8	1.2	748	1 MEPA RAT	Q64230 rattus norv
31	8	1.2	789	1 ARNT HUMAN	P27540 homo sapien
32	8	1.2	790	1 ARNT RABIT	Q02748 oryctolagus
33	8	1.2	791	1 ARNT MOUSE	P53762 mus musculu
34	8	1.2	800	1 ARNT RAT	P41739 rattus norv
35	8	1.2	837	1 ATS4 HUMAN	O75173 homo sapien
36	8	1.2	860	1 MUTS PASMU	P57972 pasteurella
37	8	1.2	884	1 SVA XYLFA	Q9ph22 xylella fas
38	8	1.2	884	1 SVA XYLFT	O87f43 xylella fas
39	8	1.2	928	1 CHS2 EXODE	P30601 exophiala d
40	8	1.2	955	1 GCSP BRAJA	Q89186 bradyrhizob
41	8	1.2	993	1 TSH DROME	P22265 drosophila
42	8	1.2	1003	1 MBD6 HUMAN	Q96dn6 homo sapien
43	8	1.2	1164	1 CNA2 HUMAN	Q9h723 homo sapien
44	8	1.2	3969	1 HRX HUMAN	Q03164 homo sapien
45	7	1.1	39	1 PRT1 BUFJA	P24641 bufo japoni

ALIGNMENTS

RESULT 1
VST2_HEVBU
ID VST2_HEVBU STANDARD; PRT; 660 AA.
AC P29326;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Burma) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31767;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
full-length viral genome.";
RL Virology 185:120-131(1991).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M73218; AAA45736.1; -
DR PIR; C40778; VHWWH2.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
Signal.
KW Q41643 volvox cart
FT SIGNAL
FT CHAIN 20 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70978 MW; 5832A013CCC4A61C CRC64;

Alignment Scores:

Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x VST2_HEVBU (1-660)

```
QY 1 ATCGCGCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20

QY 61 CCGCCCGGTGACCGGTGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGGTGCT 120
DB 21 ProProGlyGlnProSerGlyArgArgGlyArgGlyArgGlyGlyGlyGly 40

QY 121 TTCTGGGGTGACCGGGTGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60

QY 181 CCCTCGCCCGCCGATGTCACCGTGGCGCGCGGGGTGGACCTGCTGTGCCAACCGGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80

QY 241 CGACCACTCGGCTCCGCTGGCGTGACCGCGCGCGCGCGCGCGCGCGCGCTGCTCAGCTCGT 300
DB 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100

QY 301 AGACCTACCAACAGTGGCGCGCGCGCTAACCGCGGTGCGCTCGCGGCCCATGACACCCCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120

QY 361 CCGAGTCCGCTGATGTCACCTCCCGCGCGCCCATCTTGGCGCGGAGTATACATCAACA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140

QY 421 TCTCCCTTACCTCTTCGTTGGCGCACCGGCACTAACTGCTTCTTATGCGCGCCCTCTT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160

QY 481 AGTCCGCTTTTACCTCTCAGGACGGCACCAATACCATATATATGCGCACCGAAGTCTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180

QY 541 AATATGCCAGTACCGGGTGGCGGTGCCACCAATCGTTTACCGCGCGCTGGTCCCAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200

QY 601 GCTGTCGGGTTACGCCATCTCCATCTCATTCTGGCCACAGACCCACCCCGCGACG 660
DB 201 AlaValGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220

QY 661 TCGGTTGATATGAATTCATATAACTCGACGGATGTTCTGATATTTAGTCCAGCCCGGCATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240

QY 721 GCCTCTGAGTTGTGATCCCAAGTAGAGCGCTACATATCGTAACCAAGCTCGCGCTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260

QY 781 GTCGAGACCTCTCGGGTGGCTGAGGAGGCTACCTCTGCTGTGTTATGCTTTGCATA 840
DB 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280

QY 841 CATGGCTCACTCGTAAATTCCTACTATAATACACCTTATACCGGTGCGCTCGGGCTGTG 900
DB 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300

QY 901 GACTTTGCCCTTGAGCTGAGTTGCGCAACTTACCCCGGTAAACCAATAACCGGGTCC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
```

RESULT 2

VST2_HEVBU

ID VST2_HEVBU

AC P33426;

STANDARD;

PRT;

660 AA.

DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Structural protein 2 precursor (ORF2).
 OS Hepatitis E virus (strain Pakistan) (HEV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=33774;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92115700; PubMed=1731327;
 RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
 RA Malik I.A., Iqbal M., Purcell R.H.;
 RT "Characterization of a prototype strain of hepatitis E virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
 CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
 CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
 CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M80581; AAA45727.1; -;
 DR InterPro: IPR004261; SP2.
 DR InterPro: IPR008975; Viral_cap_coat.
 DR Pfam: PF03014; SP2; 1.
 KW SIGNAL.
 FT SIGNAL 1 22 BY SIMILARITY.
 FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
 SQ SEQUENCE 660 AA; 70980 MW; 8085BC53CFB46FD3 CRC64;

 Alignment Scores:
 Pred. No.: 0 Length: 660
 Score: 408.00 Matches: 508
 Percent Similarity: 99.80% Conservative: 0
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 61.82% Indels: 0
 DB: 1 Gaps: 0

 US-09-851-410A-6_COPY_5147_7129 (1-1983) x VST2_HEVPA (1-660)
 QY 289 GCCTACGCTCGTAGACCTACACAGCTGGGGCGCGCCGCTAACCGCGTGCCTCGCGCC 348
 DB 97 AlaSerArgArgProThrThrAlaGlyAlaAlaProLeuThrAlaValaProAla 116
 QY 349 CATGACACCCGCCAGTGCCTGATGCGACTCCCGCGCGCCATCTTGGCGCGCAGTAT 408
 DB 117 HisAspThrProProValProAspValAspSerArgGlyAlaAlaLeuArgGlnIlyr 136
 QY 409 AACCTATCAACATCTCCCTTACCTCTCCGTCGCCACCGGCACTAACCTGTTCTTTAT 468
 DB 137 AsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyr 156
 QY 469 GCGGCGCCCTTAGTCGGCTTTTACCCCTTCAGGACGGACCAATACCCATATATAGGCC 528
 DB 157 AlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAla 176
 QY 529 ACGGAAGCTTCTAAATTATCCACAGTACCGGGTTGCCGGTCCACATCCGTTACCGCCGC 588
 DB 177 ThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgPro 196
 QY 589 CTGGTCCCCAATGCTGTCGGCGGTGTACGCCATCTCCATCTCATCTTGGCCACAGACACC 648
 DB 197 LeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThr 216
 QY 649 ACCACCCGACGTCCGTTGATATGAATCAATAACCTCGACGGATGTTTCGTATTTAGTC 708
 DB 217 ThrThrProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuVal 236

QY 709 CAGCCGCGCATAGCCTCTGAGCTTGATGCCAAGTGAGCGCTACACTATCTGTAACCAA 768
 DB 237 GlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGln 256
 QY 769 GGCTGCGCTCGCTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGCTTGT 828
 DB 257 GlyTyrArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuVal 276
 QY 829 ATGCTTTGCATACATAGGCTCACTCTGTAATTTCTTATCTACTAATACACCCCTATACCGGTGCC 888
 DB 277 MetLeuCyIleHisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAla 296
 QY 889 CTCGGCTCTTGAGACTTTGCCCTTGAGCTTGAGTTTCGCAACTTACCCCGGTAAACACC 948
 DB 297 LeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThr 316
 QY 949 AATACGCGGCTCTCCGTTATTCAGCACTGTCGCCACCGCTTCGTCCGGTGGCGAC 1008
 DB 317 AsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAsp 336
 QY 1009 GGGACTGCCGAGCTCACACACGCGTGTCTACCCGCTTTATGAGGACCTCTATTTACT 1068
 DB 337 GlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThr 356
 QY 1069 AGTACTAATGTTGCTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCAACCTTGCT 1128
 DB 357 SerThrAsnGlyValGlyGluLeuGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAla 376
 QY 1129 GACACTCTGCTTGGCGGCTGCCGACAGAAATGATTTGCTCGGCTGGTGGCGAGCTGTT 1188
 DB 377 AspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPhe 396
 QY 1189 TACTCCGCTCCGTTGCTCAGCCCATGCGGACCGGACTGTTAAGTTGTATACATCTGTA 1248
 DB 397 TyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerVal 416
 QY 1249 GAGAAATGCTCAGCAGGATAAGGCTATTGCAATCCCGCATGACATTGACCTCGGAGATCT 1308
 DB 417 GluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSer 436
 QY 1309 CGTGTGGTTATTTCAGGATTATGATAACCAACATGAACAGATCGCGCGGCTCTTCCA 1368
 DB 437 ArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPro 456
 QY 1369 GCGCATCGCGCCCTTCTCTGCTCGAGCTAATGATGCTTGGCTCTCTCTCACCC 1428
 DB 457 AlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThr 476
 QY 1429 GCTGCGGAGTATGACCGCTCACTTATGGCTCTTCGACTGGCCGCTTATGTTTCTGAC 1488
 DB 477 AlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAsp 496
 QY 1489 TCTGTGACCTTGGTTAATGTTGCGACCGCGCGCGGCTTGGCGGCTCGCTCGATGG 1548
 DB 497 SerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyr 516
 QY 1549 ACCAAGGTCACACTTGACGGTGGCGGCTCTCCACCATCCAGCAGTACTCGAAGACCTTC 1608
 DB 517 ThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhe 536
 QY 1609 TTTGTCTCGCGCTCCGCGGTAAAGCTCTTCTCGGAGGAGGACACAACTAAAGCCGGG 1668
 DB 537 PheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGly 556
 QY 1669 TACCCCTTAATTAATAACACCACTCTGACGACCAACTGCTTGTGAGAAATGCCCGCGG 1728
 DB 557 TyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGly 576
 QY 1729 CACCGGGTCCGCTATTTCCACTTACACCACTAGCTGGGTGGTGGTCCGCTCTCCATTTCT 1788
 DB 577 HisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSer 596

```
QY 1789 GCGTTGCCGTTTTCAGCCCGCCACTCT 1815
Db 597 AlaValAlaValLeuAlaProHisSer 605
RESULT 3
VST2_HEVMY STANDARD; PRT; 660 AA.
ID VST2_HEVMY AC Q04611;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227573; PubMed=8470371;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikihisa T., Winn K.;
RT "Sequence and gene structure of the hepatitis E virus isolated from
RT Myanmar.";
RL Virus Genes 7:95-109(1993).
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10330; BAA01174.1; -.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW SIGNAL.
FT CHAIN 1 22 BY SIMILARITY.
FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70998 MW; 3A82A4EA255C6253 CRC64;
Alignment Scores:
Pred. No.: 1.3e-117 Length: 660
Score: 128.00 Matches: 128
Percent Similarity: 100.00% Conservatative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.39% Indels: 0
DB: 1 Gaps: 0
US-09-851-410a-6_COPY_5147_7129 (1-1983) x VST2_HEVMY (1-660)
QY 1 ATGCGCCCTCGGCTATTTGCTGCTCTCATGTTTGTGCTATGCTCGCCGCCCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGGTACCGCTCTGCGCGCGCTGCTGGCGGCGGCGACGCGGCTTCGGCGGTGGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly 40
QY 121 TTCGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCCGATGTCAACGCTCGCGCGGCGCTGACCTCGTGTTCGCGCAACCCGCC 240
Db 61 ProPheAlaProAspValThraAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTTGGCGTGGACGAGCCCGCCCGCGCTGCTCAGTCGT 300
Db 1789 GCGTTGCCGTTTTCAGCCCGCCACTCT 1815
Db 597 AlaValAlaValLeuAlaProHisSer 605
81 ArgProLeuGlySerAlaTTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
301 AGACCTTACACAGCTGCGCGCGCGCTAACCGCGGTCTCGTCCGCCCATGACACCCCG 360
101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
361 CCAGTGCCTGATGTCGACTCCCGC 384
121 ProValProAspValAspSerArg 128
RESULT 4
VST2_HEVRH STANDARD; PRT; 485 AA.
ID VST2_HEVRH AC Q00270;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Structural protein 2 (Fragment).
OS Hepatitis E virus (isolate Rhesus) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31766;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K.,
RA Win K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90274; BAA20910.1; -.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
FT NON_TER 1 1
FT NON_TER 485 485
SQ SEQUENCE 485 AA; 52317 MW; 5A4FD9273AC74F94 CRC64;
Alignment Scores:
Pred. No.: 1.08e-109 Length: 485
Score: 120.00 Matches: 120
Percent Similarity: 100.00% Conservatative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.18% Indels: 0
DB: 1 Gaps: 0
US-09-851-410a-6_COPY_5147_7129 (1-1983) x VST2_HEVRH (1-485)
QY 853 GTAATTCTCTATTAATACACCTATACCGGTCCCTCGGCTGTTGGACTTGCCTT 912
Db 153 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 172
QY 913 GAGCTTCAGTTTCGCAACCTTACCCCGGTAAACACCAATACGCGGGTCTCCCGTTATTC 972
Db 173 GluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTyrSer 192
QY 973 AGCACTCTCGCCACCCCTTCGTCGCGGTCCGACCGGACTGCCGAGCTCACCAACG 1032
Db 193 SerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGluLeuThrThr 212
QY 1033 GCTGCTACCGCTTTATGAGGACCTCTATTTACTAGTACTATGCTGCGGTGAGATC 1092
Db 213 AlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGlyValGlyGlu 232
```

QY 1093 GGCGGGGATAGCCCTGTTCAACCTGTTGACACTGCTGCTGGCGGCTGCCG 1152
Db 233 GlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyLeuPro 252
QY 1153 ACAGATTGATTTCGTGGCTGGTGGCAGCTGTTACTCCGTCCTCGTGTCTCAGCC 1212
Db 253 ThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAla 272

RESULT 5
VST2_HEVME
ID VST2_HEVME STANDARD; PRT; 659 AA.
AC Q03500;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor.
OS Hepatitis E virus (strain Mexico) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31768;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93079857; PubMed=1448913;
RA Huang C.-C., Nguyen D., Fernandez J., Yun K.-Y., Fry K.E.,
RA Bradley D.W., Tam A.W., Reyes G.R.;
RT "Molecular cloning and sequencing of the Mexico isolate of hepatitis
RT E virus (HEV).";
RL Virology 191:550-558 (1992).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M74506; AAA45732.1; -;
DR PIR; B44212; B44212.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW Signal.
FT SIGNAL. 1 22 BY SIMILARITY.
FT CHAIN 23 659 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 659 AA; 70640 MW; CF75E75EFD8FBE2C CRC64;
Alignment Scores:
Pred. No.: 1,03e-109 Length: 659
Score: 120.00 Matches: 120
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.18% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x VST2_HEVME (1-659)
QY 490 TPAACCCCTTCAGGAGCCACCAATACCCATATAATGGCCACGAGCTTCTAATTATGCC 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183
QY 550 CAGTACCGGGTTCGGCGTCCCAATCCCTTACCGCCCGCTGGTCCCAATCTGTCGCGC 609
Db 184 GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203
QY 610 GGTAGGCCATCTCCATCTCATCTGGCCACAGACACCACCCAGCTCCGTTGAT 669
Db 204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCAATAACCTCGACGGATGTCGTATTATTAGTCAGCCCGGCATAGCCTCTGAG 729

Db 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
QY 730 CTTGTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGCGTGGCGCTCCGTCGAGACC 789
Db 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 263
QY 790 TCTGGGGTGGCTGAGGAGGAGCTACCTGCTGGTCTTGTATGCTTTGCATACATGGCTCA 849
Db 264 SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283

RESULT 6
VST1_HEVBU
ID VST1_HEVBU STANDARD; PRT; 123 AA.
AC P29325;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 1 (ORF3).
OS Hepatitis E virus (strain Burma) (HEV), and
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31767, 33774;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Burma;
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Pakistan;
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563 (1992).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M73218; AAA45735.1; -;
DR EMBL; M60581; AAA45726.1; -;
DR PIR; B40778; VHWVHE.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
SQ SEQUENCE 123 AA; 12676 MW; 8A5A798B1B74EDE5 CRC64;
Alignment Scores:
Pred. No.: 9.66e-99 Length: 123
Score: 109.00 Matches: 109
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.52% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x VST1_HEVBU (1-123)
QY 2 TCGCGCCTCGGCCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCCGCGCCAC 61
Db 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGCCCGGTTCAGCGCTCGCGCGCTGTCGCGCGCGCGCGCGCGCGCGCGGT 121

DE Eukaryotic translation initiation factor 3 subunit 3 (eIF-3 gamma)
DE (eIF3 p38 subunit) (eIF3h).
CN TIEF3H1 OR AT1G10840 OR T16B5.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eusids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=21125776; PubMed=11042177;
RX Burks E.A., Bezerra P.P., Le H., Gallie D.R., Browning K.S.;
RT "Plant initiation factor 3 subunit composition resembles mammalian
RL J. Biol. Chem. 276:2122-2131(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Alcafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Krenetska I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman M.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vayaberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RL thaliana".
RN Nature 408:816-820(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed=14593172;
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RL genome".
RL Science 302:842-846(2003).
CC -!- FUNCTION: BINDS TO THE 40S RIBOSOME AND PROMOTES THE BINDING OF
CC METHIONYL-TRNAI AND MRNA. ASSOCIATES WITH THE P170 SUBUNIT OF
CC EIF3.
CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits.
CC -!- SIMILARITY: Contains 1 MPN (JAB/Mov34) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF285833; AAC53614.1; -

DR EMBL; AC007354; AAD31329.1; -
DR EMBL; AY054641; AAK96832.1; -
DR EMBL; AY081546; RAM10108.1; -
DR FIR; B66242; B66242.
DR InterPro; IPR003639; Pept_M67_Mov34.
DR InterPro; IPR000555; Peptidase_M67.
DR Pfam; PF01398; Mov34; 1.
DR ProDom; PD363422; Mov34_1; 1.
DR SMART; SM00232; JAB_MPN; 1.
KW Initiation factor; Protein biosynthesis.
FT CONFLICT 146 146 V -> L (IN REF. 1).
SQ SEQUENCE 337 AA; 38372 MW; F9965DA546AC977C CRC64;

Alignment Scores:
Pred. No.: 4.89 Length: 337
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.39% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x IF33_ARATH (1-337)

QY 1953 CTTAAGGCGCTGAAGCTCAGCGACAGT 1927
Db 149 LeuLysAlaLeuLysLeuSerAspSer 157
|||||
|||

RESULT 10
RL7_CLOPE STANDARD; PRT; 121 AA.
ID RL7_CLOPE
AC Q8XER7;
DC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12.
GN RPLL OR CPE2414.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater".
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -!- FUNCTION: Seems to be the binding site for several of the factors
CC involved in protein synthesis and appears to be essential for
CC accurate translation (By similarity).
CC -!- SIMILARITY: Belongs to the L12p family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF003194; BAB82120.1; -
DR HAMAP; MF_00368; -; 1.
DR InterPro; IPR008932; Ribos_L12/7_olig.
DR InterPro; IPR000206; Ribosomal_L12.
DR Pfam; PF00542; Ribosomal_L12; 1.
DR ProDom; PD001326; Ribosomal_L12; 1.
DR TIGRFAMs; TIGR00855; L12; 1.
DR Ribosomal protein; Complete proteome.
SQ SEQUENCE 121 AA; 12535 MW; 5AF9F83D2C94AC3E CRC64;

Alignment Scores:

Pred. No.: 55-2 Length: 121
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x RL7_CLOPE (1-121)

QY 83 GCCGTCGTCGGCGCGCGAGCGCG 106

Db 38 AlaValValGlyGlyAlaAlaAla 45

RESULT 11

YD2M_HERAU
ID YD2M_HERAU STANDARD; PRT; 138 AA.

AC P25279;

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 15.4 kDa protein in HgiDIIM S' region (ORF15).

OS Herpetosiphon aurantiacus (Herpetosiphon giganteus).

OC Bacteria; Chloroflexi; Herpetosiphonales; Herpetosiphonaceae;

OC Herpetosiphon.

OC NCBI_TaxID=65;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HPA2;

RX MEDLINE=92039068; PubMed=1937045;

RA Duesterhoeft A., Kroege M.;

RT "Cloning, sequence and characterization of m5C-methyltransferase-

RT encoding gene, hgiDIIM (GRCGAC), from Herpetosiphon giganteus strain

RT Hpa2.";

RL Gene 106:87-92(1991).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X55141; CAA38940.1; --

DR PIR; J0593; J0593.

KN Hypothetical protein; Restriction system; DNA-binding.

FT DNA_BIND 17 38 H-T-H MOTIF (POTENTIAL).

SQ SEQUENCE 138 AA; 15396 MW; E9CC70A0E11785A2 CRC64;

Alignment Scores:

Pred. No.: 54.1 Length: 138
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x YD2M_HERAU (1-138)

QY 822 ACCAGAGGTAGCTCTCTCTCAGC 799

Db 43 ThrArgGlySerLeuLeuLeuSer 50

RESULT 12

LED1_THEME

ID LED1_THEME STANDARD; PRT; 166 AA.

AC Q3WYCB;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE 3-isopropylmalate dehydratase small subunit 1 (EC 4.2.1.33)

DE (Isopropylmalate isomerase 1) (Alpha-IPM isomerase 1) (IPMI 1).

GN LEUD1 OR TM0292.

OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OC NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
RX MEDLINE=9287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Raft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -|- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
CC and 3-isopropylmalate, via the formation of 2-isopropylmaleate.
CC -|- CATALYTIC ACTIVITY: 3-isopropylmaleate = 2-isopropylmaleate +
CC H(2)O.
CC -|- CATALYTIC ACTIVITY: 2-isopropylmaleate + H(2)O = 2-
CC isopropylmaleate.
CC -|- PATHWAY: Leucine biosynthesis; second step.
CC -|- SUBUNIT: Heterodimer of leuC and leuD (By similarity).
CC -|- SIMILARITY: Belongs to the leuD family. LeuD 2 subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AE001711; AAD35380.1; --
DR PIR; D72394; D72394.
DR TIGR; TM0292; --
DR HAMAP; MF 01032; -- 1.
DR InterPro; IPR000573; Aconitase_C.
DR Pfam; PF00694; Aconitase_C; 1.
KW Leucine biosynthesis; Lyase; Complete proteome.
SQ SEQUENCE 166 AA; 18734 MW; D8ACF340205111A4 CRC64;

Alignment Scores:
Pred. No.: 52.7 Length: 166
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x LED1_THEME (1-166)

QY 1728 GCACCGGTCGCTATTTCACCTTA 1751
|||||
Db 19 AlaProGlyArgTyrPheHisLeu 26

RESULT 13

PSBQ_CHLRE
ID PSBQ_CHLRE STANDARD; PRT; 199 AA.
AC P12852;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oxygen-evolving enhancer protein 3, chloroplast precursor (OE33).
GN PSBQ.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OC NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=137C / CC-125;

RA Mayfield S.P., Schirmer-Rahire G., Frank H., Zuber H., Roehaia J.-D.;
RT "Analysis of the genes of the OEE1 and OEE3 proteins of the
RL Plant Mol. Biol. 12:683-693(1989).
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
CC with the photosystem II complex.
CC -1- SIMILARITY: Belongs to the psbQ family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X13832; CAA32061.1; -;
DR PIR: S05509; S05509.
DR InterPro: IPR008797; PsbQ.
DR Pfam: PF05757; PsbQ; 1.
KW Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
KW Thylakoid; Membrane.
FT TRANSIT 1 51 CHLOROPLAST.
FT CHAIN 52 199 OXYGEN-EVOLVING ENHANCER PROTEIN 3.
SQ SEQUENCE 199 AA; 21824 MW; 8CCF077334CF79D6 CRC64;

Alignment Scores:
Pred. No.: 51.3 Length: 199
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1993) x PSBQ_CHLRE (1-199)

QY 277 CGCCCGCGCTGCCTCAGTCCT 300
Db 10 ArgProAlaValAlaSerArgArg 17

RESULT 14
PSBQ_VOLCA STANDARD; PRT; 202 AA.
AC Q41643;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oxygen-evolving enhancer protein 3, chloroplast precursor (OEE3).
GN PsbQ.
OS Volvox carteri.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OX Volvocaceae; Volvox.
OX NCBI_TaxID=3067;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L. Nagariensis / HK10;
RX MEDLINE=96376499; PubMed=8761179;
RA Choi G., Przybylska M., Straus D.;
RT "Three abundant germ line-specific transcripts in Volvox carteri
RL encode photosynthetic proteins.";
RL Curr. Genet. 30:347-355(1996).
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
CC with the photosystem II complex.
CC -1- SIMILARITY: Belongs to the psbQ family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL: U22330; BAB40980.1; -;
DR PIR: S72224; S72224.
DR InterPro: IPR008797; PsbQ.
DR Pfam: PF05757; PsbQ; 1.
DR TIGRFAMs: TIGR01409; TAT signal seq; 1.
KW Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
KW Thylakoid; Membrane.
FT TRANSIT 1 54 CHLOROPLAST (POTENTIAL).
FT CHAIN 55 202 OXYGEN-EVOLVING ENHANCER PROTEIN 3.
SQ SEQUENCE 202 AA; 22290 MW; 3103DDFC84CC3180 CRC64;

Alignment Scores:
Pred. No.: 51.2 Length: 202
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1993) x PSBQ_VOLCA (1-202)

QY 277 CGCCCGCGCTGCCTCAGTCCT 300
Db 10 ArgProAlaValAlaSerArgArg 17

RESULT 15
LOLB_XANCP STANDARD; PRT; 218 AA.
AC Q8PC65;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer-membrane lipoprotein lolB precursor.
GN LOLB OR XCC0870.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Parah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.P.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RL host specificities.";
RL Nature 417:459-463(2002).
CC -1- FUNCTION: Plays a critical role in the incorporation of
CC lipoproteins in the outer membrane after they are released by the
CC lolA protein (By similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (By similarity).
CC -1- SIMILARITY: Belongs to the lolB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AE012186; AM40185.1; -.
 DR HAMAP; MF_00233; -, 1.
 DR InterPro; IPR004565; LolB.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR Pfam; PF03550; LolB; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Chapterone; Outer membrane; Lipoprotein; Transport; Protein transport;
 KW Signal; Complete proteome; Palmitate.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 218 OUTER-MEMBRANE LIPOPROTEIN LolB.
 FT LIPID 21 21 N-Palmitoyl cysteine (Potential).
 FT LIPID 21 21 S-Diacylglycerol cysteine (Potential).
 SQ SEQUENCE 218 AA; 23298 MW; 25A715B827CF782C CRC64;

Alignment Scores:
 Pred. No.: 50.6 Length: 218
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.23% Indels: 0
 DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x LolB_XANCP (1-218)

QY 326 GGGCGGGCCCGCTGTGTAGGT 303
 DB ||||||||||||||||||||
 30 GlyAlaAlaProAlaValValGly 37

Search completed: August 16, 2004, 13:48:43
 Job time : 28.2638 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:39:23 ; Search time 97.0734 Seconds

(without alignments)
12890.708 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129

Perfect score: 660

Sequence: 1 ATGCGCCTCGCCCTATTTT.....GTAAACTCGGAGTTGTAG 1983

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2033934

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-O=/cpn2_1/USPTO_spool_p/US09851410/runat_16082004_125717_9638/app_query.fasta_1.4238
-DB=SPTRMBL_25 -QWMT-fastcan -SUFFIX=oli.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -SPART=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09851410 @CGN 1.1.218 @runat_16082004_125717_9638 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

- 1: sp archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mbc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	456	69.1	660	12 Q9E8G5	Q9E8G5 hepatitis e

2	403	61.1	660	12 Q9WQA0	Q9WQA0 hepatitis e
3	395	60.0	660	12 Q89468	Q89468 hepatitis e
4	362	54.8	660	12 Q9W7W9	Q9W7W9 hepatitis e
5	344	52.1	660	12 Q9WJZ5	Q9WJZ5 hepatitis e
6	284	43.0	660	12 Q81871	Q81871 hepatitis e
7	282	42.7	660	12 Q9WLL4	Q9WLL4 hepatitis e
8	275	41.7	659	12 Q8J32	Q8J32 hepatitis e
9	264	40.0	660	12 Q69411	Q69411 hepatitis e
10	259	39.2	660	12 Q69419	Q69419 hepatitis e
11	251	38.0	660	12 Q91855	Q91855 hepatitis e
12	247	37.4	605	12 Q9YWL0	Q9YWL0 hepatitis e
13	214	32.4	660	12 Q91856	Q91856 hepatitis e
14	212	32.1	248	12 Q81860	Q81860 hepatitis e
15	197	29.8	660	12 Q81878	Q81878 hepatitis e
16	174	26.4	283	12 Q9YPB1	Q9YPB1 hepatitis e
17	168	25.5	168	12 Q9WJX0	Q9WJX0 hepatitis e
18	163	24.7	525	12 Q39947	Q39947 hepatitis e
19	149	22.6	149	12 Q9W9E8	Q9W9E8 hepatitis e
20	149	22.6	150	12 Q57042	Q57042 hepatitis e
21	149	22.6	660	12 Q89885	Q89885 hepatitis e
22	144	21.8	205	12 Q9QP98	Q9QP98 hepatitis e
23	133	20.2	135	12 Q91169	Q91169 hepatitis e
24	126	19.1	227	12 Q9W827	Q9W827 hepatitis e
25	125	18.9	258	12 Q81861	Q81861 hepatitis e
26	124	18.8	227	12 Q56048	Q56048 hepatitis e
27	124	18.8	227	12 Q56049	Q56049 hepatitis e
28	122	18.5	436	12 Q9W148	Q9W148 hepatitis e
29	119	18.0	137	12 Q8B6J4	Q8B6J4 hepatitis e
30	119	18.0	137	12 Q8B6J3	Q8B6J3 hepatitis e
31	119	18.0	137	12 Q8B6J2	Q8B6J2 hepatitis e
32	119	18.0	137	12 Q8B6J1	Q8B6J1 hepatitis e
33	119	18.0	137	12 Q8B6J0	Q8B6J0 hepatitis e
34	119	18.0	137	12 Q8B6I9	Q8B6I9 hepatitis e
35	119	18.0	137	12 Q8B6I8	Q8B6I8 hepatitis e
36	119	18.0	137	12 Q8B6I7	Q8B6I7 hepatitis e
37	119	18.0	137	12 Q8B6I6	Q8B6I6 hepatitis e
38	119	18.0	137	12 Q8B6I5	Q8B6I5 hepatitis e
39	119	18.0	137	12 Q8B6I4	Q8B6I4 hepatitis e
40	119	18.0	137	12 Q8B6I3	Q8B6I3 hepatitis e
41	119	18.0	137	12 Q8B6I2	Q8B6I2 hepatitis e
42	119	18.0	137	12 Q8B6I1	Q8B6I1 hepatitis e
43	119	18.0	137	12 Q8B6I0	Q8B6I0 hepatitis e
44	119	18.0	137	12 Q8B6H9	Q8B6H9 hepatitis e
45	119	18.0	137	12 Q8B6H8	Q8B6H8 hepatitis e

ALIGNMENTS

RESULT 1
Q9E8G5 PRELIMINARY; PRT; 660 AA.
ID Q9E8G5
AC Q9E8G5
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Abb-2B;
RX MEDLINE=20271579; PubMed=10813471;
RA van Cuyck-Gandre H., Zhang H.Y., Tsarev S.A., Warren R.L.,
RA Caudill J.D., Snellings N.J., Begot L., Innis B.L., Longer C.F.;
RT "Phylogenetically distinct hepatitis E viruses in Pakistan.";
RL Am. J. Trop. Med. Hyg. 62:187-189 (2000).
DR EMBL; AF185822; AAG16766.1;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.

SQ	SEQUENCE	660 AA; 70903 MW; 1F506BE3CFB3BACE CRC64;	
Alignment Scores:			
Pred. No.:	0	Length: 660	
Score:	456.00	Matches: 556	
Percent Similarity:	99.82%	Conservative: 0	
Best Local Similarity:	99.82%	Mismatches: 1	
Query Match:	69.09%	Indels: 0	
DB:	12	Gaps: 0	
US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q9B8G5 (1-660)			
QY	310	ACAGCTGGGCGCGCGCTAAACCGGGTGGCTCCGGCCCATGACACCCCGCAGTGCCT	369
Db	104	ThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrProValPro	123
QY	370	GATGTGCGATCCCGCGCGCATCTTGGCGCGCAGTATAACCTATCAACATCTCCCTT	429
Db	124	AspValAspSerArgGlyAlaAlaLeuArgGlnTyrAsnLeuSerThrSerProLeu	143
QY	430	ACTCTTCCGTGGCCACCGGCACTAACCTGGTCTTTATGCCGCCCTCTTAGTCCGCTT	489
Db	144	ThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeuSerProLeu	163
QY	490	TTACCCCTTCAGACGCGCACCAATACCATATATATGCGCCGAGCTTCTAATTATGCC	549
Db	164	LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla	183
QY	550	CAGTACCGGGTGGCGCGTGCACAAATCCGTTACCGCCCGCTGGTCCCAATGCTGTGGC	609
Db	184	GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly	203
QY	610	GGTTACGCCATCTCCATCTCATCTGCGCAGACACACACCGCCGAGTCCGTTGAT	669
Db	204	GlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrThrThrThrThrThr	223
QY	670	ATGAATTCATACCTCGAGGATGTCGATTTTGTAGTCCAGCCCGCATAGCTCTGAG	729
Db	224	MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu	243
QY	730	CTTGATGCCAAGTGGCGCTACACTATCGTAACCAAGCTGGCGCTCCGTCGAGACC	789
Db	244	LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThr	263
QY	790	TCTGGGTGGTGGAGGAGGCTACCTCTGGTCTGTATTATGCTTTCATACATGGCTCA	849
Db	264	SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer	283
QY	850	CTCGTAAATTCCTATACATACACCTATACCGGTGCGCTCGGCTGTGGACTTTGCC	909
Db	284	ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla	303
QY	910	CTTGAGCTTGGTTCGCAACCTTACCCCGGTAACACCAATACCGGGTCTCCCGTTAT	969
Db	304	LeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTyr	323
QY	970	TCCAGCATGCTCGCCACCGCTTCGTCGGGTGGAGCGGACTCGGAGCTCACCACC	1029
Db	324	SerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGluLeuThrThr	343
QY	1030	ACGGCTGTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGTGTGGTGAG	1089
Db	344	ThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGlyValGlyGlu	363
QY	1090	ATCGCGCGCGGATAGCCCTCACCTGTTCACCTGTGTGACACTCTCTGGCGGCGCTG	1149
Db	364	IleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeu	383
QY	1150	CGCAGAGATTGATTCTCGTGGCTGGTGGCGAGCTGTTCTACTCCGCTCCGCTCTCTCA	1209
Db	384	ProThrGluLeuLeuSerSerAlaGlyGlnLeuPheTyrSerArgProValValSer	403
QY	1210	GCCATGGCGAGCCGACTGTTAAGTGTATACATCTGTAGAGAAATGCTCAGCAGGATAAG	1269

Db	404	AlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnAspLys	423
QY	1270	GGTATTGCATCCCGCATGACATTCGAGATTCGCTGCTGCTTATTCAGGATTAT	1329
Db	424	GlyIleAlaIleProHisAspLeuGlyGluSerArgValValIleGlnAspTyr	443
QY	1330	GATAACCAACATGAACAGATCGCGCAGCCCTTCTCCAGCCCATCGCGCTTCTCT	1389
Db	444	AspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSer	463
QY	1390	GTCTTCGAGCTAATGATGTCTTGGCTCTCTCACCGCTCCGAGTATGACAGTCC	1449
Db	464	ValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyrAspGlnSer	483
QY	1450	ACTTATGGCTCTTCGACTCGCCAGTTTATGTTCTGACTCTGCTGACCTTGGTAAATGTT	1509
Db	484	ThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnVal	503
QY	1510	GCACCGCGCGCGCAGCCGCTTGCCTCGCTCGATTGGACCAAGGTCACATTGACGGT	1569
Db	504	AlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrThrLysValThrLeuAspGly	523
QY	1570	CGCCCCCTCTCCACCATCCAGAGTACTCGAAGACCTTCTTCTGCTGCGCTCCGCGGT	1629
Db	524	ArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuProLeuArgGly	543
QY	1630	AAGCTCTCTTTCGGGAGGCGAGCACAACTAAAGCCGGGTACCTTATATATTAACACC	1689
Db	544	LysLeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThr	563
QY	1690	ACTGCTAGCCACCAACTGCTTGTGAGAAATGCCCGCGGACCGGGTCGCTATTTCCACT	1749
Db	564	ThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAlaIleSerThr	583
QY	1750	TACACCATAGCTGGTGGTGGTCTCCGCTTCCATTCCTCGGTTGCGGTTTTCAGCCCC	1809
Db	584	TyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaValLeuAlaPro	603
QY	1810	CACCTCGCTGCTAGCATTCGAGGATACCTTGACTACCTGCGCTCCCGCGCCCATACTTTT	1869
Db	604	HisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAlaHisThrPhe	623
QY	1870	GATGATTTCTCCAGAGTGGCGCCCTTGGCTTGGCTTCCAGGCTCGCTTCCAGTCTACT	1929
Db	624	AspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPheGlnSerThr	643
QY	1930	GTCGCTGAGCTTCAGCCCTTAAGATGAAGGTGGGTAAACTCGGAGTTG	1980
Db	644	ValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu	660

RESULT 2

Q9WQAO	PRELIMINARY;	PRT; 660 AA.
ID	Q9WQAO	
AC	Q9WQAO;	
DT	01-NOV-1999 (TrEMBLrel. 12, Created)	
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	ORF-2.	
OS	Hepatitis E virus.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage;	
OC	Hepatitis E-like viruses.	
OC	NCBI_TaxID=12461;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=AKL-90;	
RA	MEDLINE=99350000; PubMed=10423137;	
RA	Arankalle V.A., Paranjape S., Emerson S.U., Purcell R.H.,	
RA	Walimbe A.M.;	
RT	"Phylogenetic analysis of hepatitis E virus isolates from India (1976-	
RT	1993).", Virol. 80:1691-1700(1999).	
RL	J. Gen. AF124407; AAD45493.1; -.	
DR	EMBL; AF124407; AAD45493.1; -.	

DR GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR004261; SP2.
 DR Pfam; PF03014; SP2; 1.
 DR PIR; P03014; SP2; 1.
 SQ SEQUENCE 660 AA; 71041 MW; 03B72DDF0AB7B521 CRC64;

Alignment Scores:

Pred. No.: 0 Length: 660
 Score: 403.00 Matches: 503
 Percent Similarity: 99.80% Conservative: 0
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 61.06% Indels: 0
 DB: 12 Gaps: 0

US-09-851-410A-6_COPY 5147_7129 (1-1983) x Q9WQA0 (1-660)

QY 289 GCCTCAGTCGCTAGACCTACACAGCTGGCGCGCCGCTAACCGCGTCTCGTCGGCC 348
 Db |||||
 QY 97 AlaSerArgArgProThrAlaGlyAlaAlaProLeuThrAlaValaProAla 116
 Db |||||
 QY 349 CATGACACCCGCCAGTCTGATGTCGACTCCCGCGCGCCATCTTGGCGCGGAGTAT 408
 Db |||||
 QY 117 HisAspThrProProValProAspValAspSerArgGlyAlaAlaLeuArgGlnTyr 136
 Db |||||
 QY 409 AACCTATCAATCTCCCTTACCTCTCCGTCGCGCACCGGCACTAACCTGTTCTTTAT 468
 Db |||||
 QY 137 AsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyr 156
 Db |||||
 QY 469 GCGCGCCCTCTAGTCGGCTTTTACCCCTTCAGGACGGCACCAATACCATATATGCCC 528
 Db |||||
 QY 157 AlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisMetAla 176
 Db |||||
 QY 529 ACGGAAGCTTCTAATATTATGCCAGTACCGGGTGGCGGCACCAATCGTTACCGCGCG 588
 Db |||||
 QY 177 ThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgPro 196
 Db |||||
 QY 589 CTGGTCCCAATGCTGTCGGGGTTCAGCCATCTCCATCTCAATCTCGGCACAGACCCACC 648
 Db |||||
 QY 197 LeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThr 216
 Db |||||
 QY 649 ACCACCCGAGTCTGATGATGATTAATCAATCAATCTCGACGGATGTCGATTTAGTC 708
 Db |||||
 QY 217 ThrThrProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuVal 236
 Db |||||
 QY 709 CAGCCCGCATAGCTCTGAGCTTCTGATCCCAAGTCAGCGCTACATCTCGTAACCAA 768
 Db |||||
 QY 237 GlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGln 256
 Db |||||
 QY 769 GGCTGGCGCTCGTCGAGACCTCTGGGTGGCTGAGGAGGCTACCTCTGCTTGT 828
 Db |||||
 QY 257 GlyTyrArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuVal 276
 Db |||||
 QY 829 ATGCTTTTCATACATGGCTCTGTAATCTTACTATACATACATACATACATACATACAT 888
 Db |||||
 QY 277 MetLeuGlyIleHisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAla 296
 Db |||||
 QY 889 CTCGGGCTGTGGACTTTGGCTTCTGAGCTTCTGAGTTCTGCACTTACCCCGGTAAACACC 948
 Db |||||
 QY 297 LeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThr 316
 Db |||||
 QY 949 AATACGGGGTCTCCGTTATTCAGCATCTGCTGCCACCGCTTCTGTCGGGTGGCGAC 1008
 Db |||||
 QY 317 AsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAsp 336
 Db |||||
 QY 1009 GGGACTCGGAGCTCACACACGCTCTGACCCGCTTATGAAGACCTCTATTTTACT 1068
 Db |||||
 QY 337 GlyThrAlaGluLeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThr 356
 Db |||||
 QY 1069 AGTACTAATGTGTGGTGCAGATCGCGCGGATAGCCCTTCAACCTGCTTCAACCTGCT 1128
 Db |||||
 QY 357 SerThrAsnGlyValGlyGlyIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAla 376
 Db |||||
 QY 1129 GACACTCTGTGGCGGCTCTGCCAGACAGATTTGATTTCTGCTGGGTGGCGGCTGTC 1188
 Db |||||

Db |||||
 QY 377 AspThrLeuLeuGlyGlyLeuProThrGluLeuLeuSerSerAlaGlyGlyGlnLeuPhe 396
 QY |||||
 QY 1189 TACTCCCGCTCCGCTTCTCAGCAATGCGAGCCGACTGTTAAGTTGTATACATCTGTA 1248
 Db |||||
 Db 397 TyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerVal 416
 QY |||||
 QY 1249 GAGATGCTCAGCAGATAGGTTATTCGAATCCCGCATGACATTCCTCGAGATCT 1308
 Db |||||
 Db 417 GluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSer 436
 QY |||||
 QY 1309 CCGTGGTATTATCAGGATTATGATACCAACATGAACAAGATCGCGCCGCTCTCTCCA 1368
 Db |||||
 Db 437 ArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPro 456
 QY |||||
 QY 1369 GCCCATCGCGCCCTTCT 1428
 Db |||||
 Db 457 AlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThr 476
 QY |||||
 QY 1429 GCTGCGGAGTATGACCGTCCACTTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1488
 Db |||||
 Db 477 AlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAsp 496
 QY |||||
 QY 1489 TCTGTGACCTTGGTAAATGTTGCGACCGCGCGCGAGCGCTTGGCGGTCGCTCGATTGG 1548
 Db |||||
 Db 497 SerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrp 516
 QY |||||
 QY 1549 ACCAAGTTCACACTTGCAGGTGCGCCCTCTCTCCATCCAGCAGTACTCGAAGACCTTC 1608
 Db |||||
 Db 517 ThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhe 536
 QY |||||
 QY 1609 TTTGTCCTCGCTCGCGGTAAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1668
 Db |||||
 Db 537 PheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGly 556
 QY |||||
 QY 1669 TACCCCTTAAATTAACACCACTGCTAGCGACCACTGCTTGTGAGAAATCGCGCGGG 1728
 Db |||||
 Db 557 TyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGly 576
 QY |||||
 QY 1729 CACCGGTGCTGATTTCCACTTACCACTAGCTGGGTGGTGGTGGTGGTGGTGGTGGT 1788
 Db |||||
 Db 577 HisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSer 596
 QY |||||
 QY 1789 GCGGTTGCGGTT 1800
 Db |||||
 Db 597 AlaValAlaVal 600
 RESULT 3
 Q89468 PRELIMINARY; PRT; 660 AA.
 ID Q89468;
 AC Q89468;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE UNNAMED protein product.
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=12461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yin S., Purcell R.H., Emerson S.U.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K52-87;
 RX MEDLINE=95176571; PubMed=7871758;
 RA Yin S., Purcell R.H., Emerson S.U.;
 RT "A new Chinese isolate of hepatitis E virus: comparison with strains
 recovered from different geographical regions.";
 RL Virus Genes 9:23-32(1994).
 DR EMBL; L25547; AAA91080.1; -;
 DR EMBL; L25595; AAA65490.1; -;

DR	GO; CO:0005198; F:structural molecule activity; IEA.
DR	InterPro; IPR004261; SP2.
DR	InterPro; IPR008975; Viral_cap_coat.
DR	Pfam; PF03014; SP2; 1.
DR	GO SEQUENCE 660 AA; 70379 MW; DC68116DCD639175 CRC64;
DR	GO SEQUENCE 660 AA; 70379 MW; DC68116DCD639175 CRC64;

DR GO: 0005198; F: structural molecule activity; IEA.
 DR InterPro: IPR004261; SP2.
 DR InterPro: IPR008975; Viral_cap_coat.
 DR Pfam: PF03014; SP2; 1.
 SQ SEQUENCE 660 AA; 70943 MW; 9719D30CCD932950 CRC64;

Alignment Scores:

Pred. No.: 0 Length: 660
 Score: 362.00 Matches: 562
 Percent Similarity: 99.65% Conservatives: 0
 Best Local Similarity: 99.65% Mismatches: 2
 Query Match: 54.85% Indels: 0
 DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q9W7M9 (1-660)

QY 289 GCCTCAGTCGTAGACCTACCAAGCTGGGGCGCGCGCTAACCGCGTCTGCTCGGCC 348
 DB 97 AlaSerArgArgProThrThrAlaGlyAlaAlaProLeuThrAlaValaProAla 116
 QY 349 CATGACACCCGCCAGTGCCTGATGTCGACTCCCGCGCGCCATCTTGGCGCGGAGTAT 408
 DB 117 HisAspThrProProValProAspValAspSerArgGlyAlaLeuLeuArgGlnTyr 136
 QY 409 AACCTATCAACATCTCCCTTACCTCTTCCGTGGCCACCGGCACTAACCTGGTTCTTTAT 468
 DB 137 AsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyr 156
 QY 469 GCCGCGCTCTTAGTCCGCTTTTACCCCTTAGAGCGGACCAATACCATATATGCGCC 528
 DB 157 AlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAla 176
 QY 529 ACGGAAGCTTCTTAATTATGCCAGTACCGGGTTGCCGTCACCAATCCGTTACCGCCG 588
 DB 177 ThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgPro 196
 QY 589 CTGGTCCCCAATGCTGTGGCGGTGTACGCCATCTCCATCTCTTGGCCACAGACCCACC 648
 DB 197 LeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThr 216
 QY 649 ACCACCGGACGTCGTTGATATGAAATCAATCACTCGACGGATGTCGTTATTTAGTC 708
 DB 217 ThrThrProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuVal 236
 QY 709 CAGCCCGGCATAGCTCTGAGCTGTGTATCCCAAGTGAGCGCTACCTATCGTAACCAA 768
 DB 237 GlnProGlyIleAlaSerGlnLeuValIleProSerGluArgLeuHisTyrArgAsnGln 256
 QY 769 GGCTGGCGCTCCGTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGCTTGT 828
 DB 257 GlyTyrArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuVal 276
 QY 829 ATGCTTTGCATACATGGCTCACTCGTAATTCCTATACCTATACACCTATACCGGTGCC 888
 DB 277 MetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAla 296
 QY 889 CTCGGGCTGTGGACTTTGGCCCTTGCAGCTTGAGTTTCCCACTTACCCCGGTAAACACC 948
 DB 297 LeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThr 316
 QY 949 AATACGGCGGTCTCCCGTTATTCAGCACTGTCTGCCACCGCTTCTGTCGCGGTGGGAC 1008
 DB 317 AsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAsp 336
 QY 1009 GGGACTGGCGAGCTCACCAACCGGTGCTACCCGCTTTATGAAGACCTCTATTTTACT 1068
 DB 337 GlyThrAlaGluLeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheSer 356
 QY 1069 AGTACTAATGCTGCTGGTGAAGTACCGCGCGGATAGCCCTACCCCTGTTCAACCTGCT 1128
 DB 357 SerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAla 376
 QY 1129 GACACTCTGCTTGGCGGCTGCGGACAGAAATTGATTCTGCTGGTGGTGGCGAGCTGTC 1188

DB 377 AspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPhe 396
 QY 1189 TACTCCCGTCCGTTGCTCTACGCCAATGCGAGCCGACTGTGAAGTTGATACATCTGA 1248
 DB 397 TyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerVal 416
 QY 1249 GAGATGCTCAGCAGGATAAGGTATTGCAATCCCGCATGACATTACCTCGGAGATCT 1308
 DB 417 GluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspLysLeuGlyGluSer 436
 QY 1309 CGTGTGGTTATTACAGGATTATGATACCAACATGACAGATCGGCCGAGCTTCTCCA 1368
 DB 437 ArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPro 456
 QY 1369 GCCCATCGCGCCCTTCTCTCTCTGAGCTAATGATGTGCTTGGCTCTCTCTCAC 1428
 DB 457 AlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThr 476
 QY 1429 GCTGCGGAGTATGACCGTCCACTTATGCTCTTCGACTGGCCGCGCTTATCTTCTGAC 1488
 DB 477 AlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAsp 496
 QY 1489 TCTGTGACCTTGGTTAATGTTTCCGACCGCGCGCAGCGCTTCCCGGTCTCGATTGG 1548
 DB 497 SerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyr 516
 QY 1549 ACCAAGTTCACACTTGACGTCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTTC 1608
 DB 517 ThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhe 536
 QY 1609 TTTGTCTGCGCTCCGCGGTAAAGCTCTCTTCTCGGAGGCGGACACACTAAAGCGGG 1668
 DB 537 PheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGly 556
 QY 1669 TACCTCTTAATTAACACCACTCTAGCAGCACTGCTTGTGAGAAATGCGCGCGG 1728
 DB 557 TyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGly 576
 QY 1729 CACCGGTGCTGTTATTTCCACTTACACCTAGCTGGGTGGTGGTCCGCTCTCCATTCT 1788
 DB 577 HisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSer 596
 QY 1789 GGGTGTGCGTTTAGCCCCCCTCTGCTAGCATTCGCTGAGGATACCTTGACTAC 1848
 DB 597 AlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyr 616
 QY 1849 CCGTCCCGCGCCCATCTTTTGTGATGATTTCTGCCAGAGTGGCGCCCTTGGCTTCAG 1908
 DB 617 ProAlaArgAlaHisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGln 636
 QY 1909 GGCTCGCTTTCCAGTCTACTGTCGCTGAGCTTACAGCGCTTAAGATGAAGTGGGTAAA 1968
 DB 637 GlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLys 656
 QY 1969 ACTCGGAGGTG 1980
 DB 657 ThrArgGluLeu 660

RESULT 5
 Q9WJZ5
 ID Q9WJZ5 PRELIMINARY; PRT; 344 AA.
 AC Q9WJZ5;
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Structural protein 2 (fragment).
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=12461;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=2B Abbottabad;
RA Van Cuyck-Gandre H., Clements N.J., Zhang H.Y., Caudill J.C.,
RA Cohen S.G., Coursaget P., Buisson Y., Warren R.L., Longer C.F.;
RT "Partial Sequence of HEV Isolates from North Africa and Pakistan;
RT Comparison with Known HEV Sequences.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40044; AAD09445.1; -;
DR GO; GO:0005198; F.structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR Pfam; PF03014; SP2; 1.
FT NON TER 1 1
SQ SEQUENCE 344 AA; 37146 MW; D7F3D2D9BB44098B CRC64;

Alignment Scores:
Pred. No.: 0 Length: 344
Score: 344.00 Matches: 344
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.12% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q9WJZ5 (1-344)

QY 949 AATACGCGGCTCTCCGGTTATTTCAGACACTGCTCGCCACCGCTTCGTCGCGTGGGAC 1008
DB 1 AsnThrArgValSerArgTyrSerThrAlaArgHisArgLeuArgArgGlyAlaAsp 20

QY 1009 GGGACTCGCGAGTCCACCACCGGCTGCTACCGGCTTATGAGGACCTCTATTACT 1068
DB 21 GlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThr 40

QY 1069 AGTACTAATGCTCGGTGAGTCGCGCGGATAGCCCTCAACCTGTTCAACCTGCT 1128
DB 41 SerThrAsnGlyValGlyGluIleGlyArgGlyLeuAlaLeuThrLeuPheAsnLeuAla 60

QY 1129 GACACTCTGCTCGCGGCTCCGCGACAGAAATGATTTCGTCGCGTGGTCCGAGCTGTC 1188
DB 61 AspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPhe 80

QY 1189 TACTCCGCTCCGTTGCTCAGCCATGGCGAGCGAGCTGTTAAGTTGTATACATCTGTA 1248
DB 81 TyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerVal 100

QY 1249 GACAATGCTCAGCAGGATAGGTTATGCAATCCGATGACATGACCTCGCAGAGATCT 1308
DB 101 GluAsnAlaGlnGlnAspLysGlyLeuAlaIleProHisAspLeuGlyGluSer 120

QY 1309 CGTGTGGTTATTCAGGATTATGATACCAACATGAAAGATCGCGCGCTTCTCCA 1368
DB 121 ArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPro 140

QY 1369 GCCCATCGCGGCTTCTCTGCTCTCGAGCTAATGATGCTTGGCTCTCTCTCACC 1428
DB 141 AlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThr 160

QY 1429 GCTGCGGAGTATGACACAGTCCACTTATGCTCTTCGACTCGCCGAGTTATGTTCTGAC 1488
DB 161 AlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAsp 180

QY 1489 TCTGTGACCTTGGTTAATGTTGCGACCGCGCGCGCCGTTGCCGCTCGCTCGATTGG 1548
DB 181 SerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrp 200

QY 1549 ACCAAGGTACACTGAGCGTCCGCCCTCTCCACCATCCAGCAGTACTCGAGACCTTC 1608
DB 201 ThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhe 220

QY 1609 TTTGTCTCGCGTCCCGGTAAGCTCTCTTTCTGGGAGGAGCAGCAACAATAAGCGCGG 1668
DB 221 PheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGly 240

QY 1669 TACCTTTAATTAACACCACTGTGCTAGCGACCAACTGCTTGTGAGATGCGCGCGG 1728
DB 1669 TACCTTTAATTAACACCACTGTGCTAGCGACCAACTGCTTGTGAGATGCGCGCGG 1728

DB 241 TyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGly 260
QY 1729 CACCGGCTGCTATTTCACCTTACACCTAGCTAGCTGCGTGGTCCGCTCTCCATTCT 1788
DB 261 HisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSer 280
QY 1789 GCGGTTGCGGCTTTAGCCGCCCACTCTGCGCTAGCATTCCTTGAGGATACCTTGGACTAC 1848
DB 281 AlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuGluAspThrLeuAspTyr 300
QY 1849 CTGCGCGCGCGCCATCTTTTGTGATGATTCGCTGAGAGTGGCGGCCCTTGGCCTTCAG 1908
DB 301 ProAlaArgAlaHisThrPheAspPheCysProGluCysArgProLeuGlyLeuGln 320
QY 1909 GCGTGGCTTCCAGTCTACTGCTGCTGAGCTTCAGCGCTTACGAGTGAAGTGGGTAAA 1968
DB 321 GlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLys 340
QY 1969 ACTCGGAGGTTG 1980
DB 341 ThrArgGluLeu 344

RESULT 6
ID Q81871 PRELIMINARY; PRT; 660 AA.
AC Q81871;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF 2 precursor.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP MEDLINE=92024067; PubMed=1926770;
RX Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Kikihisa T., Mizuno K., Win K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92335008; PubMed=1630924;
RA Aye T.T., Uchida T., Ma X.Z., Iida F., Shikata T., Zhuang H.,
RA Win K.M.;
RT "Complete nucleotide sequence of a hepatitis E virus isolated from the
RT Xinjiang epidemic (1986-1988) of China.";
RL Nucleic Acids Res. 20:3512-3512(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92217462; PubMed=1589964;
RA Fry K.E., Tam A.W., Smith M.W., Kim J.P., Luk K.-C., Young L.M.,
RA Platak M., Feldman R.A., Yun K.Y., Purdy M.A., McCaustland K.A.,
RA Bradley D.W., Reyes G.R.;
RT "Hepatitis E virus (HEV): Strain variation in the nonstructural gene
RT region encoding consensus motifs for an RNA-dependent RNA polymerase
RT and an ATP/GTP binding site.";

```

RL Virus Genes 6:173-185(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93348763; PubMed=8346669;
RA Bi S-L., Purdy M.A., McCausland K.A., Margolis H.S., Bradley D.W.;
RT "The sequence of hepatitis E virus isolated directly from a single
RL source during an outbreak in China.";
RL Virus Res. 28:233-247(1993).
DR EMBL; I08816; AAA03191.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2..
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 660 AA; 70979 MW; B58F233955FDD6614 CRC64;

RESULT 7
Q9WLL4
ID Q9WLL4 PRELIMINARY; PRT; 660 AA.
AC Q9WLL4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TK15/92;
RX MEDLINE=99049628; PubMed=9833882;
RA Gouvea V., Snellings N., Popek M.J., Longer C.F., Innis B.L.;
RT "Hepatitis E virus: complete genome sequence and phylogenetic analysis
RT of a Nepali isolate.";
RL Virus Res. 57:21-26(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TK15/92;
RA Gouvea V.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF051830; AAC97188.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70798 MW; A04C0185ACC085DB CRC64;

Alignment Scores:
Pred. No.: 3.63e-272 Length: 660
Score: 282.00 Matches: 282
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 42.73% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q9WLL4 (1-660)
QY 853 GTAATTCCTATCTATACACCTTACCGGTGCCCTCGGCTCTTGCACCTTCCCTT 912
DB 285 ValAsnSerTyThrAsnThrProTyThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 304
QY 913 GAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTCTCCCGTTATTC 972
DB 305 GluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTyrSer 324
QY 973 AGCAATTCCTATCTATACACCTTACCGGTGCCCTCGGCTCTTGCACCTTCCCTT 1032
DB 325 SerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGluLeuThrThr 344
QY 1033 GCTGTACCCGCTTATGAAGGACCTCTATTTTACTAGTACTAATGGTTCGGTGAGATC 1092
DB 345 AlaAlaThrArgPheMetLysAspLeuTyThrProSerAlaProSerArgProPheSerVal 464
QY 1093 GCGCGGGGATAGCCCTCACCTGTTTCAACCTTGCTGACACTCTGTTGGCGGCTGCGG 1152
DB 365 GlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuPro 384
QY 1153 ACAGAAATTCCTATCTATACACCTTACCGGTGCCCTCGGCTCTTGCACCTTCCCTT 1212
DB 385 ThrGluLeuIleSerAlaGlyGlyGlnLeuPheTyThrSerThrAsnGlyValGlyGluLe 404
QY 1213 AATGCGAGCCGACTGTGTTAAGTTGTATACATCTGTAGAGAATGCTCAGCAGGATTAAGGT 1272
DB 405 AsnGlyGluProThrValLysLeuTyThrSerValGluAsnAlaGlnGlnAspLysGly 424
QY 1273 ATTGCAATCCCGCATGACATTCACCTCGAGAACTCTCGTGTGGTTATTCAGGATTATGAT 1332
DB 425 IleAlaIleProHisAspIleAspLeuGlyGluSerArgValIleGlnAspTyrAsp 444
QY 1333 AACCAACATGACAGATCGCGAGCTTCTCCAGCCCATCTCGGCCCTTCTCTCTC 1392
DB 445 AsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerVal 464
QY 1393 CTTTCAGCTAATGATGTGTTGGCTCTCTCTCACCGCTCGCGAGTATGACCAAGTCCACT 1452
DB 1452
```

```
Db      365 GlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuPro 384
QY      1153 ACAGATTGATTCGTGGCTGGTGGCGAGCTGTTCTACTCCGTCGGTTCCTCAGCC 1212
Db      385 ThrGluLeuIleSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAla 404
QY      1213 AATGGCGAGCCGACTCTTAAGTTGTATACATCTGTAGAAATGCTCAGCAGGATAAGGGT 1272
Db      405 AsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnAspLysGly 424
QY      1273 ATTGCAATCCCGCATGACATTCACCTCGGAGAAATCTCGTGTGTTATTTCAGGATTATGAT 1332
Db      425 IleAlaIleProHisaspIleAspLeuGlyGluSerArgValValIleGlnAspTyrAsp 444
QY      1333 AACCAACATGAACAAGATCGGCGAGCCGCTTCTCCAGCCCATCGGCCCTTCTCTGTC 1392
Db      445 AsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerVal 464
QY      1393 CTTCGAGCTAATGATGTGCTTTCGCTCTCTCTCACCGCTCGCGAGTATGACCACTCCACT 1452
Db      465 LeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThr 484
QY      1453 TATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACTCTGTGCTTGTGTTAATGTGG 1512
Db      485 TyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnValAla 504
QY      1513 ACCGGCGCGAGCCGCTGTCGGTTCGCTCTCTCTCACCGCTCGCGAGTATGACCACTCCACT 1572
Db      505 ThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArg 524
QY      1573 CCCTCTCCACCATCAGCAGTAGTACTCGAAGACCTTCTTCTCTCGCGCTCCGCGGTAAG 1632
Db      525 ProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuProLeuArgGlyLys 544
QY      1633 CTCTCTTCTTCGGAGGCGAGGCACAACTAAAGCCGGGTACCTTTATATATTATACACCACT 1692
Db      545 LeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThr 564
QY      1693 GCTAGC 1698
Db      565 AlaSer 566
```

RESULT 8

```
Q8UTJ32 PRELIMINARY; PRT; 659 AA.
ID Q8UTJ32 AC Q8UTJ32
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22072932; PubMed=12076829;
RA Jameel S., Zafrullah M., Chawla Y.K., Dilawari J.B.;
RT "Reevaluation of a North India isolate of hepatitis E virus based on
RT the full-length genomic sequence obtained following long RT-PCR.";
RL Virus Res. 86:53-58 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Jameel S., Zafrullah M., Chawla Y.K., Dilawari J.B.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF459438; AAM66330.1;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 659 AA; 70842 MW; ABCC177EBE6012C8 CRC64;
```

```
Alignment Scores: 3.58e-265 Length: 659
Pred. No.: 275.00 Matches: 375
Percent Similarity: 99.73% Conservative: 0
Best Local Similarity: 99.73% Mismatches: 1
Query Match: 41.67% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q8UTJ32 (1-659)
QY 853 GTAAATCTTACTATAATACACCCCTATACGGTGCCTCGGCTGTGTGGACTTTGCCCTT 912
Db 284 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 303
QY 913 GAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGGTCTCCCTTTATTC 972
Db 304 GluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTyrSer 323
QY 973 AGCACTGCTCGCCACCGCTTGTTCGGTGGCGAGCGGAGCTCCGAGCTCACCAACG 1032
Db 324 SerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGluLeuThrThr 343
QY 1033 GCTGCTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGTTGCGTGGATC 1092
Db 344 AlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGlyValGlyGluIle 363
QY 1093 GCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTCTGACACTCTGTGCGGCGCTGCCG 1152
Db 364 GlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuPro 383
QY 1153 ACAGAAATTGATTCGTGGCTGGTGGCCAGCTGTTCTACTCCGTCGGTTCCTCAGCC 1212
Db 384 ThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAla 403
QY 1213 AATGGCGAGCCGACTCTTAAGTTGTATACATCTGTAGAGAAATGCTCAGCAGGATAAGGGT 1272
Db 404 AsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnAspLysGly 423
QY 1273 ATTGCAATCCCGCATGACATTCGCTCGGAGAAATCTCGTGTGTTATTCAGGATTATGAT 1332
Db 424 IleAlaIleProHisaspIleAspLeuGlyGluSerArgValValIleGlnAspTyrAsp 443
QY 1333 AACCAACATGAACAAGATCGGCGAGCCCTTCTCCAGCCCATCGGCCCTTCTCTGTC 1392
Db 444 AsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerVal 463
QY 1393 CTTCGAGCTAATGATGTGCTTTCGCTCTCTCTCACCGCTCCGAGTATGACCACTCCACT 1452
Db 464 LeuArgAlaAsnAspValLeuTrpLeuSerLeuThrGlyAlaGluTyrAspGlnSerThr 483
QY 1453 TATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACTCTGTGTTAATGTTGCG 1512
Db 484 TyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnValAla 503
QY 1513 ACCGGCGCGAGCCGCTTTCGGTTCGCTCGATTGGACCAAGTCACTTACGCTCGC 1572
Db 504 ThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArg 523
QY 1573 CCCTCTCCACCATCAGCAGTAGTACTCGAAGACCTTCTTGTCTCGCGCTCCGCGGTAAG 1632
Db 524 ProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuProLeuArgGlyLys 543
QY 1633 CTCTCTTCTTCGGAGGCGAGGCACAACTAAAGCCGGGTACCTTTATATTATAACACCACT 1692
Db 544 LeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThr 563
QY 1693 GCTAGCACCACACTGCTTGTTCGAGAAATGCCCGCGGACACCGGTCGCTATTTCACATAC 1752
Db 564 AlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyr 583
QY 1753 ACCACTAGCTGGGTGCTGTCCTGCTCTCATTTCTCGCGTTCGCTTTAGCCCCCCCAC 1812
Db 584 ThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaValLeuAlaProHis 603
```

```

QY 1813 TCTGGCTAGCATTCCTTGGAGATACCTTGACTACCTGCGCGCCCATCTTTTGTAT 1872
Db 604 SerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAlaHisThrPheAsp 623
QY 1873 GATTTCCTCCAGAGTGGCGCCCTTGGCTTCCAGGCTGGCTTCCAGTCTACTGTC 1932
Db 624 AspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrVal 643
QY 1933 GCTGAGCTTCAGCGCTTAAGATGAAGTGAGTGAAGCTGAACTCGGAGTTG 1980
Db 644 AlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 659

RESULT 9
Q69411
ID Q69411 PRELIMINARY; PRT; 660 AA.
AC Q69411;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=hev037;
RA Donati M.C., Fagan E.A., Harrison T.J.;
RT "Sequence analysis of full length HEV clones derived directly from
human liver in fulminant hepatitis E.";
RL (In) Rizzetto M., Purcell R.H., Gerin J.L., Verme G (eds.);
RL VIRAL HEPATITIS AND LIVER DISEASE, pp.313-316,
RL Edizioni Minerva Medica, Torino (1997).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=hev037;
RA Harrison T.J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X98292; CAA66937.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71024 MW; F37E3FEFF7A4ED1 CRC64;

Alignment Scores:
Pred. No.: 3 52e-254 Length: 660
Score: 264.00 Matches: 264
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.00% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q69411 (1-660)

QY 919 GAGTTTCGCAACCTTACCCCGGTACACCAATACGCGGTCTCCGTTATTCAGCACT 978
Db 307 GluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThr 326
QY 979 GCTCGCCACGCTTCGTGCGGTGGCGGAGTCCGAGTCCACCAACGCGTGTCT 1038
Db 327 AlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAla 346
QY 1039 ACCCGCTTATGAGGACCTCTATTTTACTAGTACTAATGGTCTCGGTGAGATCGCGCG 1098
Db 347 ThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGlyValGlyGluLeuGlyArg 366
QY 1099 GGGATAGCCCTCACCTGTTTCAACCTTGTGACACTCTGCTGCGCGCTGCCGACGAA 1158
Db 367 GlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGlu 386
QY 1159 TTGATTCGTCGCTGGTGGCCAGCTGTTTCTACTCCCGTCCCGTGTGCTCAGCCAAATGGC 1218

```

```

Db 387 LeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGly 406
QY 1219 GAGCGGACTGTTAGTTGTATACATCTGTAGAGAAATGCTCAGCAGATAAAGGTATTGCA 1278
Db 407 GluProThrValLysLeuTyrThrSerValGluAsnAlaGlnAspLysGlyIleAla 426
QY 1279 ATCCCGCATGACATTGACCTCGGAGAAATCTCGTGGTGTATTTCAGGATTATGATACCAA 1338
Db 427 IleProHisAspIleAspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGln 446
QY 1339 CATGAACAAGATCGCGCGAGCCCTTCTCCAGCCCCCATCGCGCCCTTCTGTCTCTCGA 1398
Db 447 HisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeuArg 466
QY 1399 GCTAATGATGCTTTGGCTCTCTCTCACCCTGCCGAGTAGTACCACTTATATGGC 1458
Db 467 AlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGly 486
QY 1459 TCTTCGACTGCGCCAGTTTATGTTCTGACTCTGTGACCTTGGTTAATGTTGCGACCGGC 1518
Db 487 SerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnValAlaThrGly 506
QY 1519 GCGAGCGCGTTCGCCGCTCGCTCGATTGGACCAAGGTTCACACTTGCAGTTCGCCCTC 1578
Db 507 AlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeu 526
QY 1579 TCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTCGCGCTCCGCGGTAAGCTCTCT 1638
Db 527 SerThrIleGlnTyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSer 546
QY 1639 TTCTGGGAGCAGCAGCAACTAAGCCGCTACCTTATATATATACACCACTGCTGAGC 1698
Db 547 PheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSer 566
QY 1699 GACCACTGCTT 1710
Db 567 AspGlnLeuLeu 570

RESULT 10
Q69419
ID Q69419 PRELIMINARY; PRT; 660 AA.
AC Q69419;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF1, ORF2 & ORF3.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RA von Brunn A., Seebach J., Thyagarajan S.P., Mohanavalli B., Menon T.,
RA Froesner G.;
RT "PCR amplification, cloning and sequence determination of a hepatitis
E virus isolate from Madras, India.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X99441; CAA67804.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70960 MW; DAC3DF95F91689F8 CRC64;

Alignment Scores:
Pred. No.: 3 49e-249 Length: 660
Score: 259.00 Matches: 559
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 3
Query Match: 39.24% Indels: 0
DB: 12 Gaps: 0

```

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q69419 (1-660)	
QY	295 CGTCGTAGACTACACAGCTGGGGCGCGCCGCTAAACCGCGTTCGCTCCGGCCCATGAC 354
Db	99 ArgArgArgProThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAsp 118
QY	355 ACCCGCGAGTGCCTGATGCTGACTCCGGCGGCGCCATCTGCGCGCGAGTATAACCTA 414
Db	119 ThrProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeu 138
QY	415 TCACATCTCCCTTACCTCTCCGTGGCCACCGGCACTAACCTGGTCTTTATGCGGCC 474
Db	139 SerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAla 158
QY	475 CCTCTTAGTCCGCTTTTACCCCTTCAGGACGGCCACCAATACCCCATATAATGGCCACGGAA 534
Db	159 ProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGlu 178
QY	535 GCTTCTAATTATGCCAGTACCGGGTTCGCGTCCGACCAATCCGTACCGCCCGCTGGTC 594
Db	179 AlaSerAsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuVal 198
QY	595 CCCAATGCTGCGCGGTACGCCATCTCCATCTCATTTCTGGCCACAGACACCAACACACC 654
Db	199 ProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThr 218
QY	655 CCGACGTCGCTGATATGAATCAATAACCTCGACGGAGTTCGTATTTTATGTCAGGCC 714
Db	219 ProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnPro 238
QY	715 GGCATAGCTCTGAGCTTGATCCCAAGTCAGCGCTACACTATCGTAACCAAGGCTGG 774
Db	239 GlyIleAlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyr 258
QY	775 CGCTCCGTCGAGACTCTCGGGTGGCTCAGAGGAGGCTACCTCTGTGCTTGTATGCTT 834
Db	259 ArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeu 278
QY	835 TGCATACATGGCTCACTCGTAATTCCTATACCTATACACCTATACCGGTGCTCCGG 894
Db	279 CysIleHisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGly 298
QY	895 CTGTTGGACTTTGCCCTTGAGCTTGAGTTTGCACACTTACCCCGGTAAACACCAATACG 954
Db	299 LeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThr 318
QY	955 CGGTCTCCGCTTATTCAGACACTGCTCGCCACCGCTTCGTGCGGTGCGGACGGGACT 1014
Db	319 ArgValSerA-gTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThr 338
QY	1015 GCGGAGCTCACACACCGGTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACT 1074
Db	339 AlaGluLeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThr 358
QY	1075 AATGCTGTCGTGAGATCGCGCGGAGTAGCCCTCACCTGCTTCAACTCTGCTCACACT 1134
Db	359 AsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThr 378
QY	1135 CTGCTTGGCGGCTCCCGACAGAAATTGATTTCTCGTGGTGGGCGAGCTGTTACTCC 1194
Db	379 LeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSer 398
QY	1195 GTCCTCGTGTCTACCCATCGGACCGGACTGTTAGTTGTTATACATCTCTAGAGAT 1254
Db	399 ArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrArgSerValGluAsn 418
QY	1255 GCTCAGCAGATAAGGATTCCAATCCCGCATGACATTGACCTCGGAGAACTCGTGTG 1314
Db	419 AlaGlnGlnAspGlyGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgVal 438
QY	1315 GTTATTTCAGATATATGATACCAACATGAACAGATCGCGCGAGCTTCTCCAGCCCA 1374
Db	439 ValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaPro 458

QY	1375 TCGCGCCCTTCTCTGCTCTCGAGCTAATGATGCTGCTTTGGCTCTCTCTCACCGCTGCC 1434
Db	459 SerArgProPheSerAlaLeuArgAlaAsnAspValLeuThrLeuSerLeuThrAlaAla 478
QY	1435 GAGTATGACAGTCCACTTATGCTCTTCGACTGGCCCGCACTTATGTTTCTGACTCTGTG 1494
Db	479 GluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerVal 498
QY	1495 ACCTTGTTAATGTTGCGACCGCGCGCAGCGCTGCGCGCTCGATTGGACCAAG 1554
Db	499 ThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrThrLys 518
QY	1555 GTCACACTTACCGCTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTC 1614
Db	519 ValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheVal 538
QY	1615 CTGCGCTCCGCGTAAAGCTCTCTTTCTGGAGGACGACCACTAAAGCCGGGTACCT 1674
Db	539 LeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrPro 558
QY	1675 TATTAATTATACACCACTGCTAGCGCACTGCTTCGAGATGCGCGGCGCACCGG 1734
Db	559 TyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArg 578
QY	1735 GTCGCTATTTCACACTTACACACTAGCTGCTGGTGTCTGCTCCGCTCTCCATTTCTCGGTT 1794
Db	579 ValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaVal 598
QY	1795 GCCGTTTTAGCCCCCACTCTGCGCTAGCTAGCTTGTGAGGATACCTTGGACTACCTGCC 1854
Db	599 AlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAla 618
QY	1855 CGCGCCCATACTTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCTTCCAGGCTGC 1914
Db	619 ArgAlaHisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCys 638
QY	1915 GCTTTTCAGTCTACTGTCGCTGAGCTTCAGCGCTTAAAGTGAAGTGGTAAACTCGG 1974
Db	639 AlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArg 658
QY	1975 GAGTTG 1980
Db	659 GluLeu 660

RESULT 11

091855	PRELIMINARY;	PRT;	660 AA.
ID	091855;		
AC	091855;		
DT	01-NOV-1998 (TrEMBLrel. 08, Created)		
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Structural protein.		
OS	Hepatitis E virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage;		
OC	Hepatitis E-like viruses.		
OX	NCBI_TaxID=12461;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=93-Egypt;		
RX	MEDLINE=99105430; PubMed=9890424;		
RA	Tsarev S.A., Binn L.N., Gomatos P.J., Arthur R.R., Monier M.K.,		
RA	van Cuyck-Gandre H., Longer C.F., Innis B.L.;		
RT	"Phylogenetic analysis of hepatitis E virus isolates from Egypt.;"		
RL	J. Med. Virol. 57:68-74 (1999).		
DR	EMBL; AF051351; AAC35761.1; -.		
DR	GO; GO:0005198; F:structural molecule activity; IEA.		
DR	InterPro; IPR004261; SP2.		
DR	InterPro; IPR008975; Viral_cap_coat.		
DR	Pfam; PF03014; SP2; 1.		
SQ	SEQUENCE 660 AA; 71040 MW; 044FF5EA7C492791 CRC64;		

Alignment Scores:

FT NON TER 605 605
SQ SEQUENCE 605 AA; 64740 MW; A088F76F874B2B72 CRC64;

Alignment Scores:
Pred. No.: 3,46e-237 Length: 605
Score: 247.00 Matches: 247
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 37.42% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q9YWL0 (1-605)

QY 1075 AATGGTCTCGTAGATCGCGCGGATAGCCCTCACCTGTTCACCTTGCTGACACT 1134
DB 359 AsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThr 378
QY 1135 CTGCTTGGCGGCTCGCCGACAGAAATGATTTCGTCGGCTGGTGGCCAGCTGTTCTACTCC 1194
DB 379 LeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSer 398
QY 1195 CGTCCCGTTCTCAGCAATGGCGAGCGACCTGTTAAGTTGTATACATCTAGAGAAT 1254
DB 399 ArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsn 418
QY 1255 GCTCAGCAGGATGAAGGTATTGCAATCCGCGATGACATTGACCTCGGAGAAATCTCGGTG 1314
DB 419 AlaGlnGlnAspLysGlyIleAlaProHisAspLeuGlyGluSerArgVal 438
QY 1315 GTTATTCAGGATTATGATAACCAACATGAACAAGATCGCGCGAGCTTCTCCAGCCCA 1374
DB 439 ValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaPro 458
QY 1375 TCGCGCCCTTCTCTGTCCTCGAGCTAATGATGCTGTTGGCTCTCTCTCACCGCTGCC 1434
DB 459 SerArgProHisSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAla 478
QY 1435 GAGTATGACCACTTATGCTCTTCGACTGGCCGAGTTTATGTTTCTGACTCTGTG 1494
DB 479 GluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerVal 498
QY 1495 ACTTGTGTTAATGTTGGACCGCGCGAGCGCTTCCCGGTGCTCGATTGGACCAAG 1554
DB 499 ThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrThrLys 518
QY 1555 GTCACACTTCACGTCGCCCCCTCTCCACCATCCAGCAGTACTCGAGACCTTCTTGTTC 1614
DB 519 ValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheVal 538
QY 1615 CTGCGCGCTCCGCGTAAAGCTCTCTTCTGGGAGCGAGGCAACAATAAGCCGGGTACCCCT 1674
DB 539 LeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrPro 558
QY 1675 TATAATTATTAACCACTGTAGCGACCAACTGTTGTGAGAAATGCGCGCGGCAACCGG 1734
DB 559 TyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGlyHisArg 578
QY 1735 GTCGCTATTTCCACTTACCACTAGCTGGGTGCTGGTCCGCTCTCCATTTCTGGGTT 1794
DB 579 ValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSerAlaVal 598
QY 1795 GCCGTTTTAGCCCCCACTCT 1815
DB 599 AlaValLeuAlaProHisSer 605

RESULT 13
O91856 PRELIMINARY; PRT; 660 AA.
AC O91856;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.

OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=94-Egypt;
EX MEDLINE=99105430; PubMed=9890424;
RA Tsarev S.A., Binn L.N., Gomatos P.J., Arthur R.R., Monier M.K.,
van Cuyck-Gandre H., Longer C.F., Innis B.L.;
RT "Phylogenetic analysis of hepatitis E virus isolates from Egypt.";
J. Med. Virol. 57:68-74(1999).
DR EMBL: AF0511352; AAC35764.1; -
DR GO: GO:0005398; P: structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71026 MW; 16C560FA16941F2A CRC64;

Alignment Scores:
Pred. No.: 3,23e-204 Length: 660
Score: 214.00 Matches: 214
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 32.42% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x O91856 (1-660)

QY 1069 AGTACTAATGGTCTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTTGCT 1128
DB 357 SerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAla 376
QY 1129 GACACTCTGTTGGCGGCTCGCGACAGAAATGATTTCGTCGGCTGGTGGCCAGCTGTC 1188
DB 377 AspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPhe 396
QY 1189 TACTCCCGTCCCTGCTCTCAGCAATGGCGAGCGACTGTTAAGTTGTATACATCTGTA 1248
DB 397 TyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerVal 416
QY 1249 GAGAAATGCTCAGCAGGATGAAGGTATTGCAATCCCGCATGACATTGACCTCGGAGAACT 1308
DB 417 GluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspLeuLeuGlyGluSer 436
QY 1309 CGTGTGTTTATTCAGGATTATGATAACCAACATGAACAAGATCGCGCGAGCTTCTTCCA 1368
DB 437 ArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPro 456
QY 1369 GCCCCATCGCGCTTCTCTGTCCTCGAGCTAATGATGCTGTTGGCTCTCTCTCACC 1428
DB 457 AlaProSerArgProHisSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThr 476
QY 1429 GCTCGCGAGTATGACCACTTATGCTCTTCGACTGGCCGAGTTTATGTTTCTGAC 1488
DB 477 AlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAsp 496
QY 1489 TCTGTGACCTTGTTAATGTTGGACCGCGCGAGCGCTTCCCGGTGCTGCGATGTTG 1548
DB 497 SerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrp 516
QY 1549 ACCAAGTCCACCTTCAGCGTGGCCCCCTCTCCACCATCCAGCAGTACTCGAGACCTTC 1608
DB 517 ThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhe 536
QY 1609 TTTGTCCTCGCTCCGCGTAAAGCTCTCTTTCTGGGAGCGAGGCAACAATAAGCCGGG 1668
DB 537 PheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGly 556
QY 1669 TACCCTTATAATTATAACCACTGCTAGCGACCAACTGCTT 1710
DB 557 TyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeu 570

```
RESULT 14
Q81860 PRELIMINARY; PRT; 248 AA.
ID Q81860 AC Q81860;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Structural protein (fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BurmaEpi-1-3;
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BurmaEpi-1-3;
RX MEDLINE=92396046; PubMed=1387921;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Zhuang H.,
RA Win K.M.;
RT "Sequence comparison of the capsid region of Hepatitis E viruses
RT isolated from Myanmar and China.";
RL Microbiol. Immunol. 36:615-621(1992).
DR EMBL; D10333; BAA01179.1; -;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR Pfam; PF03014; SP2; 1.
FT NON_TER 1
FT NON_TER 248
SQ SEQUENCE 248 AA; 26859 MW; 3C8139211E2D045F CRC64;

Alignment Scores:
Pred. No.: 3,75E-202 Length: 248
Score: 212.00 Matches: 212
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 32.12% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q81860 (1-248)
QY 1219 GAGCCGACTGTAAGTTGTATACATCTGTAGAGAATCTCAGCAGGATTAAGGGTATTGCA 1278
DB 1 GluProThrValIysLeuTyrThrSerValGluAsnAlaGlnAspLysGlyIleAla 20

QY 1279 ATCCCGCATCACATTGACCTCGGAGATCTCGTGTGTTATTTCAGGATTATGATACCAA 1338
DB 21 IleProHisAspLeuAspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGln 40

QY 1339 CATCAACAGATCGCCGACGCTTCCTCCAGCCCATCGCCGCTTCTCTGTCCTTCGA 1398
DB 41 HisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeuArg 60

QY 1399 GCTTAATGATGCTTTGGCTCTCTCACCCTGCGCAGTATGACCATTTATGTC 1458
DB 61 AlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGly 80

QY 1459 TCTTCGAGTCCCGCAGTTTATGTTTCTGACTCTGTGACCTTGGTATATGTTGCGACCGGC 1518
DB 81 SerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnValAlaThrGly 100

QY 1519 GCGCAGGCGCTTCCCGGTCTGCTCGATTGGACCAAGGTCACTTGACGTCGCCCTC 1578
DB 101 AlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeu 120

QY 1579 TCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCCGCTCCGCGGTAAAGCTCTCT 1638

RESULT 15
Q81878 PRELIMINARY; PRT; 660 AA.
ID Q81878 AC Q81878;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Complete genome sequence.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ug179;
RA Uchida T.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; D11093; BAA01867.1; -;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70937 MW; 5412313F1A2EF4A9 CRC64;

Alignment Scores:
Pred. No.: 3,14E-187 Length: 660
Score: 197.00 Matches: 197
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.85% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q81878 (1-660)
QY 1006 GAGGAGCTCCGAGCTCACCACGCGTCTACCGCTTTATGAGGACCTCTATTT 1065
DB 336 AspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLysAspLeuTyrPhe 355

QY 1066 ACTAGTACTAATGGTGTGCGTGAGATCGCGCGGATAGCCCTCACCTGTTCAACCTT 1125
DB 356 ThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeu 375

QY 1126 GCTGACACTCTGCTTGGCGGCTTCCGACAGAAATTGATTTCTGCGGTGGTGGCCAGCTG 1185
DB 376 AlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeu 395

QY 1186 TTCTACTCCGTCGCCGTGTCAGCCAAATGGCGAGCCGACTGTTAAGTTGTATACATCT 1245
DB 396 PheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSer 415

QY 1246 GTAGAGATGCTCAGCAGGATAGGGTATTGCAATCCGCGATGACATTGACCTCGGAGAA 1305
DB 416 ValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspLeuGlyGlu 435
```

```
QY 1306 TCTCGTGTGGTTATTTCAGATTATGATAACCAACATGAACAGATCGGCCGACGCTTCT 1365
Db |||||
QY 436 SerArgValValIleGlnAspTyrAspAsnGlnHisGlnAspArgProThrProSer 455
Db |||||
QY 1366 CCAGCCCATCGGCCCTTCTCTGCTCCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTC 1425
Db |||||
QY 456 ProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeu 475
Db |||||
QY 1426 ACCGCTGCCGAGTATGACAGTCCACTTATGGCTCTTCGACTGGCCGAGTTTATGTTTCT 1485
Db |||||
QY 476 ThrAlaAlaGluTyrAspGlnSerThrTyrglySerSerThrGlyProValTyrglySer 495
Db |||||
QY 1486 GACTCTGTGACCTTGGTTAATGTTGCGACCGCGCGCAGGCGGTTGCCCGGTGCGTCGAT 1545
Db |||||
QY 496 AspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAsp 515
Db |||||
QY 1546 TGGACCAAGTTCACACTTGACGTCGCCGCTCTCCACCATCCAGCAGTAC 1596
Db |||||
QY 516 TrpThrIysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyr 532
Db |||||
```

Search completed: August 16, 2004, 13:55:37
Job time : 121.073 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:41:04 ; Search time 20.0151 Seconds
(without alignments)
10229.686 Million cell updates/sec

US-09-851-410A-6_COPY_5147_7129

Perfect score: 660
Sequence: 1 ATGGCCCTCGGCTATTTT.....GTAAACTCGGAGTTGTAG 1983

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Word size: 1

Total number of hits satisfying chosen parameters: 663654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/uspto spoop/US09851410/runat 16082004 125718 9682/app_query.fasta_1.4238
-DB=Issued Patents AA -OMT=fastan -SUFFIX=oli.rai -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptd -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09851410 @CGN 1 1 45 @runat 16082004 125718 9682 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	660	100.0	660	1	US-08-240-049B-15
2	660	100.0	660	1	US-08-259-148A-19
3	660	100.0	660	1	US-08-484-054-19
4	660	100.0	660	2	US-07-876-941A-19
5	660	100.0	660	3	US-08-478-507-8
6	660	100.0	660	3	US-08-542-634-13
7	660	100.0	660	3	US-09-128-275A-8
8	660	100.0	660	3	US-08-477-292-13
9	660	100.0	660	4	US-09-553-427-8
10	660	100.0	660	4	US-09-462-606-50
11	660	100.0	660	4	US-07-870-985A-19
12	660	100.0	660	5	PCT-US95-13703-13

13	549	83.2	549	3	US-08-542-634-15	Sequence 15, Appl
14	549	83.2	549	3	US-08-477-292-15	Sequence 15, Appl
15	549	83.2	549	5	PCT-US95-13703-15	Sequence 15, Appl
16	540	81.8	540	3	US-08-542-634-25	Sequence 25, Appl
17	540	81.8	540	5	PCT-US95-13703-25	Sequence 25, Appl
18	525	79.5	525	3	US-08-542-634-27	Sequence 27, Appl
19	525	79.5	525	5	PCT-US95-13703-27	Sequence 27, Appl
20	512	77.6	660	4	US-09-172-699-2	Sequence 2, Appl
21	436	66.1	436	1	US-08-259-148A-17	Sequence 17, Appl
22	436	66.1	436	1	US-08-484-054-17	Sequence 17, Appl
23	436	66.1	436	2	US-07-876-941A-17	Sequence 17, Appl
24	436	66.1	436	4	US-07-870-985A-17	Sequence 17, Appl
25	413	62.6	552	4	US-09-172-699-16	Sequence 16, Appl
26	413	62.6	561	4	US-09-172-699-20	Sequence 20, Appl
27	408	61.8	660	4	US-09-462-606-12	Sequence 12, Appl
28	396	60.0	660	4	US-09-462-606-49	Sequence 49, Appl
29	327	49.5	327	1	US-08-240-049B-13	Sequence 13, Appl
30	327	49.5	327	1	US-08-259-148A-15	Sequence 15, Appl
31	327	49.5	327	1	US-08-484-054-15	Sequence 15, Appl
32	327	49.5	327	2	US-07-876-941A-15	Sequence 15, Appl
33	327	49.5	327	3	US-08-542-634-17	Sequence 17, Appl
34	327	49.5	327	3	US-08-477-292-17	Sequence 17, Appl
35	327	49.5	327	4	US-07-870-985A-15	Sequence 15, Appl
36	327	49.5	327	5	PCT-US95-13703-17	Sequence 17, Appl
37	321	48.6	660	3	US-08-840-316-2	Sequence 2, Appl
38	321	48.6	660	3	US-08-809-523-2	Sequence 2, Appl
39	321	48.6	660	3	US-08-471-971-2	Sequence 2, Appl
40	321	48.6	660	4	US-09-402-776-2	Sequence 2, Appl
41	321	48.6	660	5	PCT-US93-08849A-2	Sequence 2, Appl
42	321	48.6	660	5	PCT-US93-08849-2	Sequence 2, Appl
43	284	43.0	660	4	US-09-462-606-56	Sequence 56, Appl
44	264	40.0	660	4	US-09-462-606-53	Sequence 53, Appl
45	259	39.2	660	4	US-09-462-606-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1

US-08-240-049B-15
; Sequence 15, Application US/08240049B
; Patent No. 5686239
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Tam, Albert W.
; APPLICANT: Varbough, Patrice O.
; TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,049B
; FILING DATE: 09-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles K. Sholtz
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.35; G32P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
; INDIVIDUAL ISOLATE: ORF-2
US-08-240-049B-15

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-08-240-049B-15 (1-660)

QY	1	ATCGCGCCTCGGCTATTTTGTGCTGCTCTCATGTGTTTGTCTATGCTGCGCGCCCA	60
Db	1	MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20
QY	61	CCGCCGGTCAAGCTGCGCGCGTCTGCGCGGTCGTGGCGCGCAGCGCGGTTCCGCGGT	120
Db	21	ProProGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly	40
QY	121	TTCTGGGTCAACCGGTTGATTTCTAGCCCTTCGCAATCCCTATATTCATCCCAAC	180
Db	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn	60
QY	181	CCCTCGCCCCCGATGTCACCGTGGCGCGGCGCTGGAGCTGCTGTCGCCAACCGCC	240
Db	61	ProPheAlaProAspValThrAlaAlaAlaGlyProArgValArgGlnProAla	80
QY	241	CGACCACTCGGCTCCGCTGGCTGACCGGTCACCGGCGCGCGCGCGGCTGCTGCT	300
Db	81	ArgProLeuGlySerAlaTrpArgGlnAlaGlnArgProAlaValAlaSerArg	100
QY	301	AGACCTACCAAGCTGGGGCGCGCTAACCGCGTCTCGCGCCATGACACCGCG	360
Db	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro	120
QY	361	CGAGTCCGTGATCGACTCCCGCGCGCATCTTCGCGCGGAGTATACCTATCAACA	420
Db	121	ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr	140
QY	421	TCTCCCTTACCTCTTCGGTGGCCACCGGCACTAACCTGGTTCTTTATGCGCCCTCT	480
Db	141	SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu	160
QY	481	AGTCCCGCTTTACCCCTTCAGGACGGCACCAATACCCATATATGCGCGGAGCTTCT	540
Db	161	SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer	180
QY	541	AATATGCCAGTACCGGTTGCCGTGCCAATCCGTTACCGCGCTGTTCGCCAAT	600
Db	181	AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn	200
QY	601	GCTGTGGGGTTACGCCATCTCCATCTCATCTGCGCCACAGCACACCGCGGAGC	660
Db	201	AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr	220
QY	661	TCCGTTGATGATGAATCAATCACTCGAGGATGTTTCGTATTTAGTCCAGCCGSCATA	720
Db	221	SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle	240
QY	721	GCTGTGACTGTGATCCCAAGTAGCGCCTACACTATCGTAACCAAGGCTGGCGCTCC	780
Db	241	AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer	260
QY	781	GTCCGAGACCTCTGGGTGCTGAGGAGGAGTACCTCTGCTGTTGTTATGCTTGCATA	840

Db	261	ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle	280
QY	841	CATGGCTCACTGTAATTCCTATACACCTATACCGGTGCCCTCGGGCTGTG	900
Db	281	HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu	300
QY	901	GACTTTGCCCTTGAGCTTGCAACCTTACCCCGGTACACCAATACCGGGTC	960
Db	301	AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal	320
QY	961	TCCCGTTATTCACACTGCTCGCCACCGCTTCGTCGCGGTGCGGACGGAGCTGCCGAG	1020
Db	321	SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu	340
QY	1021	CTCACCACACGCGTCTACCCGCTTATGAAGACCTCTATTTTACTAGTACTAATGGT	1080
Db	341	LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly	360
QY	1081	GTCCGTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTTGCTGACACTCTGCTT	1140
Db	361	ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu	380
QY	1141	GGCGGCTCGCGACAGAAATTGATTTCGTGCGCTGGTGGCCAGCTGTCTACTCCCGTCCC	1200
Db	381	GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro	400
QY	1201	GTGTCCTCAGCCCAATGCGGAGCGACTGTAACTGTATACATCTGTAGAGATGCTCAG	1260
Db	401	ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln	420
QY	1261	CAGGATAAGGATTTCGAATCCCGCATCGACATGACCTCGAGAATCTCGTGTGTTATT	1320
Db	421	GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle	440
QY	1321	CAGGATTATGATAACCAACATGAACAGATCGCGCGCGCTTCTCCAGCCCCCATCGCGC	1380
Db	441	GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg	460
QY	1381	CTTTCTCTGCTTCGAGTATGATGTGCTTGGCTCTCTCTCACCGCTGCGGAGTAT	1440
Db	461	ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr	480
QY	1441	GACCACTCCACTTATGGCTCTTCGACTGGCCCAAGTTATGTTTCTGACTCTGTGACCTTG	1500
Db	481	AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu	500
QY	1501	GTTAATGTTGCGACCGCGCGCGCGCTTGGCCGCTCGATTCGATTCGACCAAGGTCA	1560
Db	501	ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr	520
QY	1561	CTTGACGTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTCTGCCG	1620
Db	521	LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro	540
QY	1621	CTCCGCGGTAAAGCTCTTTCTGGGAGCGGACCAACTAAAGCCCGGTACCTTATAAT	1680
Db	541	LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn	560
QY	1681	TATAACACACTGTAGCAGCAACTGCTTGTGAGAATGCCCGCGGACCGGGTTCGT	1740
Db	561	TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla	580
QY	1741	ATTTCCACTTACCACTAGCTGGTGTGCTGCTCCGCTCCCTCCCTTCGCTGCGTCCGCT	1800
Db	581	IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal	600
QY	1801	TTAGCCCCCACTCTCGCTAGCTTGTGTGAGGATACCTTGGACTACCTCCCGCGCGC	1860
Db	601	LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla	620
QY	1861	CATACTTTTGATGATTTCTGCCACAGTGCCTCCCTCTGGCTTCAGGGCTGCGCTTTC	1920
Db	621	HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe	640

QY 1921 CAGTCTACTGCTGAGCTTCAGCGCTTAAGATGAGCTGGCTAAGAACTCGGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 2

US-08-259-148A-19
Sequence 19, Application US/08259148A
Patent No. 5741490
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R.
APPLICANT: Bradley, Daniel W.
APPLICANT: Twu, Jr-Shin
APPLICANT: Purdy, Michael A.
APPLICANT: Tam, Albert W.
APPLICANT: Krawczynski, Krzysztof Z.
APPLICANT: Yarbough, Patrice D.
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/259,148A
FILING DATE: 13-JUN-1994
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APR-1989

REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO

ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
US-08-259-148A-19

Alignment Scores: 0 Length: 660
Pred. No.:

Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-08-259-148A-19 (1-660)
QY 1 ATGGCGCCCTGGCGCTATTTTCTGCTCTCTCATGTGTTTTTGCCTATGCTCCCGCGCCA 60
Db 1 MetArgProArgProLleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGGTGACCGGCTCTGGCGCGGTGTGGGGCGGCACGCGGGTTCGGGGGTGGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlySerGlyGlyGly 40
QY 121 TTTCTGGGGTGAACCGGGTGTATCTCAGCCCTTCCCAATCCCTATATTCATCCACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTGGCCCGGCTGTCACCGCTGCGCGCGGGGTGGACCTCGTGTTCGGCAACCCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGGCTTGGCGTGACCGACCGCCGCGCGCGGTTCCTCTCAGTCGT 300
Db 81 ArgProLeuGlySerAlaTipArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACTTACACAGCTGGGGCGCGGCTTAACCGGGTTCGTCGGCGCCCATGACACCCCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTGAGCTCCCGCGCGCCCATCTTGGCGCGGCAGTATAACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCTGGCCACCGGCACTAACCTGGTTCTTTATGCGCCCTCTTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGCTTTTACCTTACGACCGCACCATATCCATATATATGATGCGCACGAAAGTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGTTCCCGTGCACAAATCCGTTACCGCGCGCTGGTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaAlaAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTGGCGGTTACGCCATCTCCATCTCATTTCTGGCCACAGACACACACCGCGAGC 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrProThr 220
QY 661 TCGGTTGATATGAATTCAATACCTTCGAGGATGTTGTTATTTAGTCAGCGCCGGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGTGATCCCAAGTAGCGCTACACTATCTGTAACCAAGCTGGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGTACTCTGGTCTTGTATGCTTTGTCATA 840
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTACTCGTAAATTCCTATATAACCTATACCCCTATACCGGTGCGCTCGGGCTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAACCAATACCGGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCCAGCACTGCTCGCCACCGCTTCTGTCGGGTGCGGACGCGGACTGCCGAG 1020

321 SerArgTyrSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
 1021 CTACACACACCGCTGCTACCGCTTATGAAGACCTCTATTTACTAGTAAAGGT 1080
 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCCCTGTTCACCTTGCTGACACTCTGCTT 1140
 361 ValGlyGluLeuGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
 1141 GCGGGCTGCGACAGAAATGATTTCTGCTGGCTGGCCAGCGTTCCTACTCCGTCGCC 1200
 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
 1201 GTTGTCTCACCAATGCGAGCGACGCTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
 1261 CAGGATAAGGGTATTGCAATCCGCAATGACCTCGAGAAATCTCGTGGGTATT 1320
 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle 440
 1321 CAGGATTATGATTAACCAATGACAGATGCGCGAGCGCTTCTCCAGCCCATCGCGC 1380
 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
 1381 CTTTCTCTCTCTCTCGAGCTAATGATGCTTTGGCTCTCTCCAGCGTGGCGAGTAT 1440
 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
 1441 GACCACTCACTTATGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
 1501 GTTAATGTTGCGACCGCGCGAGCGCTTGGCGGTGCGTTCGATTCGACCAAGGTGACA 1560
 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
 1561 CTTGACGCTGCGCCCTCTCCACATCCAGCTACTCGAAGACCTTCTTTGTCCTGCGG 1620
 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
 1621 CTCGCGGTGAAGCTCTCTTCTGGAGGAGGAGGAGCAAACTAAAGCGGGTACCTTATAAT 1680
 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
 1681 TATAACACCACTGCTAGCACCACTGCTTTCGAGAAATGCGCGCGGCGGACCGGTGCGT 1740
 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
 1741 ATTTCCACTTACACCACTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1800
 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
 1801 TTAGCCCCCACTGCTGCGCTAGCATGCTTGAGGATACCTTGAGCTACCTCGCGCGCGC 1860
 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
 1861 CATACCTTTGATGATTTCTGCCAGAGTCCCGCCCTTGGCTTCAGGCTGCGGTTC 1920
 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlyCysAlaPhe 640
 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAAGTGAAGTGGTAAAGTGGAGGTG 1980
 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 3

US-08-484-054-19
 ; Sequence 19, Application US/08484054
 ; Patent No. 5770689
 ; GENERAL INFORMATION:
 ; APPLICANT: Reyes, Gregory R.

; APPLICANT: Bradley, Daniel W.
 ; APPLICANT: Twu, Jr-Shin
 ; APPLICANT: Purdy, Michael A.
 ; APPLICANT: Tam, Albert W.
 ; APPLICANT: Krawczynski, Krzysztof Z.
 ; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dehlinger & Associates
 ; STREET: 350 Cambridge Avenue, Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,054
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 870,985
 ; FILING DATE: 20-APRIL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 822,335
 ; FILING DATE: 17-JAN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 505,888
 ; FILING DATE: 05-APRIL-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 420,921
 ; FILING DATE: 13-OCTOBER-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 367,486
 ; FILING DATE: 16-JUNE-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 336,672
 ; FILING DATE: 11-APRIL-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 208,997
 ; FILING DATE: 17-JUNE-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sholtz, Charles K.
 ; REGISTRATION NUMBER: 38,615
 ; REFERENCE/DOCKET NUMBER: 4600-0093.38
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 324-0880
 ; TELEFAX: (415) 324-0960
 ; INFORMATION FOR SEQ ID NO: 19:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 660 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
 ; US-08-484-054-19

Alignment Scores:
 Pred. No.: 0 Length: 660
 Score: 660.00 Matches: 660
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-08-484-054-19 (1-660)

QY 1 ATGCGCCCTCGGCCTATTGTTGCTCTCATGTTTTCCTATGCTGCGCGCCA 60


```

1 MetArgProArgProIleLeuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
61 CCGCCCGCTGAGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGTGGT 120
21 ProProGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlyGlyGly 40
121 TTCGGGGTGACCGGGTGTATCTCAGCCCTTCGAATCCCTATATTCATCAACCAAC 180
41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
181 CCCTTCGCCCGCGATGTCACCGCTGCGCGCGCGCTGCGCGCGCTGCGCGCGCGCG 240
61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
241 CGACCACTCGCGCTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTCAGCTGT 300
81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
301 AGACTTACCAAGCTGGCGCGCGCTAACCGCGCTGCTCCGGCGCGCGCGCGCGCGCG 360
101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
361 CCAGTGCCTGATGTCGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
421 TCTCCCTTACCTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 480
141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
481 AGTCGGCTTTACCTCTCAGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
541 AATTATGCCAGTACCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
181 AsnTyrAlaGlnTyrArgValAlaAlaAlaAlaThrIleArgTyrArgProLeuValProAsn 200
601 GCTGTCGGCGGTTACGCCATCTCCATCTCATCTCTGCGCGCGCGCGCGCGCGCGCG 660
201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
661 TCCGTTGATATGAATCAATCACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
721 GCCTCTGAGCTTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGCGCTGGCGCTCC 780
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
781 GTCGAGACCTCTGGGTGGCTGAGGAGGAGTACCTCTGGTCTGTTTANGCTTTCGATA 840
261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
841 CATGGCTCATGTAATTCCTTACTATACACCTATACCGCTGCGCTGCGCGCGCGGTG 900
281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
901 GACTTTGCGCTTGAGCTGAGTTTCCCACTTACCGCGCGCGCGCGCGCGCGCGCGCG 960
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
961 TCCCGTTATTCAGCACTGCTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
321 SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
1021 CTCACCAACCAAGCTGCTACCGCTTATCAAGACCTCTATTTTACTACTACTATGCT 1080
341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
1081 GTGGTGAGATCGCGCGCGGATAGCCCTACCTGTTCACTTGTGACACTCTGCTT 1140

```

```

361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
1141 GCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTTC 1200
381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
1201 GTTGCTCTCAGCAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 1260
401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
1261 CAGGATTAAGGATTAATCCCGCATGACATTCACCTCCGAGAAATCTCGTGTGTTATT 1320
421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
1321 CAGGATTAATGATAACCAACATGAACATCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
441 GlnAspTyrAspAsnGlnHisGluAspArgProThrProSerProAlaProSerArg 460
1381 CTTTCTCTGCTCTCGAGCTTAATGATGTGTGTTGGTCTCTCTCAGCGTGGCGAGTAT 1440
461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
1441 GACAGTCCACTTATGCTCTTCGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1500
481 AspGlnSerThrTyrGlySerSerThrThrGlyProValTyrValSerAspSerValThrLeu 500
1501 GTTAATGTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTCA 1560
501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
1561 CTTGACGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 1620
521 LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuPro 540
1621 CTCGCGGTAACTCTCTCTTGGGAGGAGCGAGCAACTAAAGCCGGTACCTTATAT 1680
541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
1681 TATAACACCACTCTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 1740
561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
1741 APTTCCACTTACCACTAGCTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
1801 TTAGCCCCCACTCTCGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
1861 CATACTTTTGTGATTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTTC 1920
621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
1921 CAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

```

RESULT 4

```

US-07-876-941A-19
; Sequence 19, Application US/07876941A
; Patent No. 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; TITLE OF INVENTION: Antibodies
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250

```

CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/876,941A
 FILING DATE: 01-MAY-1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 822,335
 FILING DATE: 17-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 505,889
 FILING DATE: 05-APRIL-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 420,921
 FILING DATE: 13-OCTOBER-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 367,486
 FILING DATE: 16-JUNE-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 336,672
 FILING DATE: 11-APRIL-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 208,997
 FILING DATE: 17-JUNE-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Sholtz, Charles K.
 REGISTRATION NUMBER: 38,615
 REFERENCE/DOCKET NUMBER: 4600-0093.33
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 660 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
 US-07-876-941A-19

Alignment Scores:
 Pred. No.: 0
 Score: 660.00
 Length: 660
 Matches: 660
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 Indels: 0
 Gaps: 2

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-07-876-941A-19 (1-660)

QY 1 ATGCGCCCTCGGCTATTTGTTGCTCTCCTCATGTTTTCCTATGCTGCGCGGCA 60
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 CGCGCCGCTCAGCCGCTCTGCGCGCCGCTGCTGCGCGCGCGAGCGGCTTCGCGCGTGT 120
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 21 ProProGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlySerGlyGly 40
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 TTCTGGGGTGACGGGTGATTTCTAGCCCTTCGCAATCCCTATATTATCCAAACCAAC 180
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyriIleHisProThrAsn 60
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 CCCTTCGCCCGGATGTACCGCTGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 240
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyProArgValArgGlnProAla 80
 QY 241 CGACCACTCGGCTCGCTGCGTGACAGGCCCGCCGCGTGGCTCCTCAGTGT 300
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 81 ArgProLeuGlySerAlaTrpArgSpGlnAlaGlnArgProAlaValAlaSerArgArg 100
 QY 301 AGACCTACCACTGCGCGCGCGCTAACCGCGGTGCTGCGCGCCCATGACACCCCG 360
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
 QY 361 CCAGTGCTGATGTGCACTCCCGCGCGCCATCTTGGCGCGGCACTATAACTATCAACA 420
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 ProValProAspValAspSerArgGlyAlaIleleuargGlnTyriAsnLeuSerThr 140
 QY 421 TCTCCCTTACCTCTTCCGTGGCCACCGGCACCTAACCTGTCTTTTATGCGCCCTCTT 480
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyriAlaAlaProLeu 160
 QY 481 AGTCCGCTTTTACCCCTTCAGGACGCGCAACATACCCATATATATGCGCAGGAGCTTCT 540
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
 QY 541 AATTATGCCCAGTACCGGGTTGCCGTGCCACCAATCCGTTACCGCCCGCTGGTCCCAAT 600
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 AsnTyriAlaGlnTyriArgValAlaArgAlaThrIleargTyriArgProLeuValProAsn 200
 QY 601 GCTGTGCGGGTTACGCCATCTCCATCTANTCTGGCCACAGACCACCAACCCCGACG 660
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 201 AlaValGlyGlyTyriAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
 QY 661 TCCGTTGATATGAATTCMAATACTCGACGATGTTCTGATTTTAGTCCAGCCCGCAT 720
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
 QY 721 GCCTCTGAGCTTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGCTCGCGCTCC 780
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyriArgAsnGlnGlyTrpArgSer 260
 QY 781 GTGAGACCTCTGGGGTGCTGAGGAGGAGGTACTCTTGGTCTTGTATTGCTTTGCATA 840
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
 QY 841 CATGCTCACTCGTAATTCCTATATAACACCTATACCGCTCGGTCCTCGGGCTGTG 900
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 281 HisGlySerLeuValAsnSerTyriThrAsnThrProTyriThrGlyAlaLeuGlyLeu 300
 QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGAAACCTTACCCCGGTAACCAACCAATACGCGGTC 960
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
 QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCTGCGGGTGCGGACGGGACTGCGGAG 1020
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 321 SerArgTyriSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
 QY 1021 CTCACCAACCAACCGCTGCTACCCGCTTATGAAGGACCTCTATTTTACTAGTACTAATG 1080
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 341 LeuThrThrThrAlaAlaThrargPheMetIysAspLeuTyriPheThrSerThrAsnGly 360
 QY 1081 GTCGGTGAGATCGCGCGGGATAGCCCTCACCTTGTTCACCTTGTCTGACACTCTGTCTT 1140
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
 QY 1141 GSCGCGCTCGCGACAGAAATGATTTGCGGGTGTGCGGCTGCTTCTATCTCCGCTGCC 1200
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyriSerArgPro 400
 QY 1201 GTTGTCTCAGCCAAATGCGGAGCCGACTGTAAAGTTGTATACATCTGTAGAGATGCTCAG 1260
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 401 ValValSerAlaAsnGlyGluProThrValIysLeuTyriThrSerValGluAsnAlaGln 420
 QY 1261 CAGGATAGGGTATTGCAATCCCGCATGACATTCACCTCGGAGATCTCGTGTGTTATT 1320
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440

QY 1321 CAGGATTATGATAACCAACATGATCGCGAGCGCTTCTCCAGCCCATCGCGC 1380
DB 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTGCCGAGTAT 1440
DB 461 ProPheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCAGCTATGCTCTTCTGAGTGGCCAGTTATGTTCTGACTCTGTGACTTGG 1500
DB 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGGCAGCGCGCGAGCGCTTGGCCGCTCGCTGATTTGGACCAAGGTCACA 1560
DB 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspThrThrValThr 520
QY 1561 CTTGACGGTGGCCCTCTCCACCTCCAGCAGTACTCGAAGACCTTTTGTGCTGCCG 1620
DB 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAAAGCTCTCTTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
DB 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGCTAGCGACCACTGCTTGTGAGAATGCGCGGCGAGCGGTCGCT 1740
DB 561 TyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTCCACTTACACCACTAGCTGGTGTGCTGCTGCCCTTCCATTTCTCGGTTGCCGTT 1800
DB 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCATCTGCGGTAGCTTGTGAGGATACCTGAGTACCTGCGCGCGCGC 1860
DB 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCTTCAGGCTGCGGCTTC 1920
DB 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGAGCTTCCAGCCCTTAAGATGAGGTGGTAAACTCGGAGTTG 1980
DB 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 5

US-08-478-507-8
; Sequence 8, Application US/08478507
; Patent No. 6120988
; GENERAL INFORMATION:
; APPLICANT: Reves, Gregory R
; APPLICANT: Yarborough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fy, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,507

FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA: US 08/279,823
FILING DATE: 25-JUL-1994
PRIOR APPLICATION DATA: US 07/681,078
FILING DATE: 05-APR-1991
PRIOR APPLICATION DATA: US 07/505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA: US 07/420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA: US 07/367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA: US 07/336,672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA: US 07/208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0183.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-507-8

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-08-478-507-8 (1-660)

QY 1 ATGCGCCCTCGGCTATTTTGTGCTCTCATGTTTTCCTATGCTGCGCGGCCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGGTCAGCCCTGTGGCGCGCTGTGGCGCGCGCGCGCGCGGTTCGCGCGTGT 120
DB 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly 40
QY 121 TTCTGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCTTCGCGCCCGATGTACCGCTCGCGCGCGGCTGGACCTCGGTTCGCGCAACCGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGGCTTGGCGTGACCAAGGCCAGCGCCCGCGCTTCGCTTCAGCTGT 300
DB 81 ArgProLeuGlySerAlaTyrArgAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACCTACACAGCTGGGGCGCGCGCTTAACCGCGGTTCGCGCGCGCTGACACCGCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTCCCTGATGTGCTCCCGCGCGCGCTTTCGCGCGCGCTATTAACCTATCAACA 420
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120

Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
Qy 421 TCTCCCTTACCTCTTCGCTGGCCACCGGACACTAACTGGTCTTTATGCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
Qy 481 AGTCCCGCTTTTACCCCTTACAGGACGGCACCAATACCCATATTAATGGCCACGGAAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
Qy 541 AATTATGCCAGTACCGGGTTGCCGTGCCACAAATCCGTTACCGCCCGCTGGTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
Qy 601 GCTGTCCGGCTTACGCCATCTCATCTCTCTGCGCCACAGACACACACCCCGGACG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
Qy 661 TCCGTTGATATGAATTCATAACCTCGACGGATGTTCTGTATTTTAGTCCAGCCCGGCATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
Qy 721 GCCTCTGAGCTTGATGCCAAGTACGAGCGCTACACTATCGTAAACCAAGCTGGCGTCC 780
Db 241 AlaSerGluLeuValIleProSerSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
Qy 781 GTCCAGACCTCGGGTGGCTGAGGAGGAGCTACCTCTGCTCTGTATGCTTTGCATA 840
Db 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
Qy 841 CATGGCTCACTCGTAAATTCCTATPACTATAACCTTATACCGGTGCCCTCGGCTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
Qy 901 GACTTTCCTTACCTTGGTGTAGTTTCGCAACCTTACCCCGGTAAACCAATACCGCGTCC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
Qy 961 TCCGTTATTCACACTGCTCGCACCGCTCGCTCGGTGGGACGGAGCTCGCGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
Qy 1021 CTCACCAACCGCTGCTACCGCTTTATGAAGACCTCTATTTACTAGTACTAATGCT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
Qy 1081 GTCCGTGAGATCGCCGGGATAGCCCTCACTCTGTCACTTGTGACACTCTGCTT 1140
Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
Qy 1141 GCGGCTGCGCACAGAAATGATTTGCTGGCTGGTGGCCAGCTGTCTACTCCCGTCCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlnLeuPheTyrSerArgPro 400
Qy 1201 GTTCTCTCAGCAATGGCGACCGACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
Qy 1261 CAGATAAGGTATGCAATCCCGCATGACATTCACCTCGGAGATCTCGTGTGTTATT 1320
Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
Qy 1321 CAGATTATGATAACCAATGAACAGATCGCGGACGCTTCTCCAGCCCGCTCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
Qy 1381 CCTTCTCTGCTTCGAGCTTAATGATGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
Qy 1441 GACCAGTCCATTATGGCTCTCGACTGGCCAGTTATGTTCTGACTCTGTGACCTTG 1500
Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500

Qy 1501 GTTAATGTTTCGACCGCGCGCGCGGTTGCCGTTGCGTTCGATTGGACCAAGTCCACA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrThrLysValThr 520
Qy 1561 CTTGAGCGTTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCGCG 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
Qy 1621 CTCGGGTAAGCTCTCTTTCTGGAGGAGGACCAACTAAAGCCGGGTACCTTATAAT 1680
Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
Qy 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTTCGAGAATCCCGCGGACCGGCTCGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
Qy 1741 ATTTCCACTTACACCACTAGCTGGTGGTGGTGGTCCGCTCTCCATTTCTGGGTGCGCTT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
Qy 1801 TTAGCCCCCAGCTCGCTAGCATTTGCTTGAAGATACCTTGGACTACCTGCGCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
Qy 1861 CATACTTTTGATGATTTCTCCAGAGTGGCGCCCTTGGCTTCAGGCTGCGCTTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
Qy 1921 CAGTCTACTCTGCTGAGCTTACCGCTTAAAGATGAAGTGGGTAAACTCGGAGCTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 6

US-08-542-634-13
; Sequence 13, Application US/08542634
; Patent No. 6214970
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; APPLICANT: McAtee, C. Patrick
; APPLICANT: Yarbough, Patrice O.
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/542,634
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```
;
; HYPOTHETICAL: NO
; ORIGINAL SOURCE: Hepatitis E Virus (Burma strain)
; INDIVIDUAL ISOLATE: ORF-2
; INDIVIDUAL ISOLATE: US-08-542-634-13
US-09-851-410a-6_COPY_5147_7129 (1-1983) x US-08-542-634-13 (1-660)

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-851-410a-6_COPY_5147_7129 (1-1983) x US-08-542-634-13 (1-660)
QY 1 ATGCGCCCTCGCCCTATTTGTTGCTGCTCCTCATGTTTTTGCCTATGCTGCGCCGCCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCCCGGTGACCCCTCTCGCCCGCGCTGCTGGCGCGCGAGCGGTTCCGGCGGTGGT 120
DB 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly 40
QY 121 TTCTGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCGATGTCAACCGCTCGCGCGGTGAGCCTCTGCTGTCGCCAACCCGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGGCTGGCGTACCAGCGCCAGCGCCCGCGCTGTCCTCAGTCTGT 300
DB 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACCTACACAGCTGGGGCGCGCGCTAACCGGGTGCCTCGCGCCCATACACACCCCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTCCGACTCCCGCGCGCCATCTTTGCGCGCGCAGTATCAACCTATCA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACTCTCTCGTGGCCACCGCCTACCTGCTGTTCTTATGCGCCGCTCTT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTTACCCCTTCAGACGCGCACCAATACCATATATATGCGCCACGAGCTTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGTTCCCGTGCCCAATCCGTTACCGCGCTGTGCCCAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaAlaAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCCGGGTTAGCCATCTCCATCTCTGTCGCCACAGACACACACCGCCGCGAG 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
QY 661 TCCGTTGATATCAATTCAATAACCTCGACGGATGTTGCTATTTAGTCCAGCGCCGCAT 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGTGATCCCAAGTAGCGCCTACACTATCGTAACCAAGCTGGCGCTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGTACCTCTGGTCTCTGTTATGTTGTCATA 840
DB 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTCGTAATTCCTATATACTAATACACCCCTATACCGGTGCCCTCGGGCTGTG 900
```

```
DB HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTCCCGCGGTAAACACCAATACGGGGTC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGTTATTTCAGCACCTGCTCGCCACCGCTTCGTCGCGTCCGAGCGGACTGCCGAG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCGCTGCTACCGCTTTATGAAGACCTCTATTATTTACTAGTACTAATGCT 1080
DB 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCTTGTTCACCTTGTCTGACACTCTGCTT 1140
DB 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCGCACAGAATTGATTTTCGTCGCGTGTGGCCAGCTGTTCTACTCCCGTCCC 1200
DB 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
DB 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGTAAGGTATTTGCAATCCCGCATGACATGACCTCGGAGAAATCTCGTGTGTTATT 1320
DB 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
QY 1321 CAGGATATTATGATTAACCAACATGAACATCGCGCGAGCCCTTCTCCAGCCCATCGCGC 1380
DB 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGCTCTCGAGCTAATGATGCTTTCCTCTCTCTCACCGCTGCCAGTAT 1440
DB 461 ProPheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACAGTCACTTATGCTTCTGACTGCGCCAGTTATGTTCTGACTCTCTGACCTTG 1500
DB 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGGACCGCGCGCAGCGCTGTCGCTCGCTCGATTGACCAAGGTCA 1560
DB 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTCTGCTGCG 1620
DB 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAGCTCTCTTCTGGAGGAGCGACCAACTAAAGCCGGGTACCTTATAT 1680
DB 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTTCGAGAATGCGCGCGGACCGGTCGCT 1740
DB 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACCACTAGCCTGCGTGGTGTGCTGCCCTTCCATTTCTCGGTTGCCGTT 1800
DB 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCGCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTGCCCGCGCC 1860
DB 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGATGATTTCTGCCACAGTGCGCCCTTGGCTTTCAGGGCTGCGCTTTC 1920
DB 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTCAGCTTCAGCGCCCTTAAGATCAAGGTGGGTAAACCTCGGAGTTG 1980
```

Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

```

RESULT 7
US-09-128-275A-8
; Sequence 8, Application US/09128275A
; Patent No. 6229005
; GENERAL INFORMATION:
; APPLICANT: Reves, Gregory R
; APPLICANT: Yarborough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6229005-A/No. 6229005-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,275A
; FILING DATE: 03-AUG-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-128-275A-8

```

Alignment Scores:
 Pred. No.: 0
 Score: 660.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 Length: 660
 Matches: 660
 Conservative: 0
 Mismatches: 0
 Indels: 0

DB:	3	Gaps:	0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-128-275A-8 (1-660)			
QY	1	ATCGCCCTCGGCTATTTTGTCTCTCTCATGTTTGTCTATGTCGCGGCCCA	60
Db	1	MetArgProArgProIleLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20
QY	61	CCGCGCGGTCAGCCGCTCGCCCGCTGCGGCGGCGGCGGCGGCTTCCGCGGTGT	120
Db	21	ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly	40
QY	121	TTCTGGGGTGACCGGCTTGATTCACGCCCTTCGCAATCCCTATATTCATCAACA	180
Db	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProIleHisProThrAsn	60
QY	181	CCCTTCGCCCCCGATGTCACCGCTGCGGCGGCGGCTGGACCTGCTGTCGCCAAC	240
Db	61	ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnPro	80
QY	241	CGACCACTCGGCTCCGCTTGGCGTGACCGGCCCGGCGGCGGCGGCTGCTCAGCT	300
Db	81	ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArg	100
QY	301	AGACCTACCACTCGGCGGCGGCTAACCGCGTCCGCTCGCGCCCATGACACCCCG	360
Db	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThr	120
QY	361	CCAGTCCCTGATGTCGACTCCCGCGGCGCATCTTCGCGGCGGAGTATACCTATCA	420
Db	121	ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTrpAsnLeuSerThr	140
QY	421	TCTCCCTTACCTCTTCGCTGCGGCGGCGGCTTAACCTGCTGCTTTATGTCGCG	480
Db	141	SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTrpAlaAlaPro	160
QY	481	AGTCCGCTTTTACCCCTTCAGGACGCGCAATACCATATATATGCGCGGAGCTTCT	540
Db	161	SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer	180
QY	541	AATTATGCCAGTACCGGCTTCCGCGGCGGCGGCTTACCGCGGCTGCTGCTCCCA	600
Db	181	AsnTyrAlaGlnTyrArgValAlaAlaAlaThrIleArgTyrArgProLeuValPro	200
QY	601	GCTGTCGCGGCTTACCGCATCTCCATCTCATCTGCGGCGGCGGCGGCGGCGGCG	660
Db	201	AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr	220
QY	661	TCCGTTGATATGAATTCATTAACCTCGAGGATGTTGCTATTTTCTAGTCCAGCG	720
Db	221	SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGly	240
QY	721	GCCTCTGAGCTTGTGATCCCAAGTACGCGCTTACACTATCGTAACCAAGCTGGCG	780
Db	241	AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer	260
QY	781	GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGGTCTGTTATGCTTGC	840
Db	261	ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle	280
QY	841	CATGCTCATCTGTAATTCCTATATAACCTATACACCTATACCGGCTCGGCTGTTG	900
Db	281	HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeu	300
QY	901	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAACCAATACGCGGTC	960
Db	301	AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal	320
QY	961	TCCGCTTATTCAGGACTGCTCGGCGGCGGCTTCTGTCGCGGTCGCGGAGTCCGCG	1020
Db	321	SerArgTyrSerSerThrAlaArgHisArgLeuA-gargGlyAlaAspGlyThrAlaGlu	340
QY	1021	CTCACCACCGGCTGCTACCCGCTTTTATGAAGGACCTCTATTTTACTAGTACTAT	1080

361 CCAGTGCCTGATGTGACCTCCGCGCGCGCATCTTGGCGCGGAGTATACCTATCAACA 420
Db |||||
121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
QY |||||
421 TCTCCGCTTACCTCTCCGCGCGCGCATCTTGGCGCGGAGTATACCTATCAACA 480
Db |||||
141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY |||||
481 AGTCCGCTTTTCCGCTTCCGCGCGCGCATCTTGGCGCGGAGTATACCTATCAACA 540
Db |||||
161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY |||||
541 AATTATGCCAGTACCGGCTTCCGCGCGCGCATCTTGGCGCGGAGTATACCTATCAACA 600
Db |||||
181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY |||||
601 GCTGTCGGGGTTCAGCATCTCCATCTCATCTGTCGCGCGCATCTTGGCGCGGAGTATACCTATCAACA 660
Db |||||
201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY |||||
661 TCCGCTGATATGAATTCATTAACCTCGAGCGATGTTCTGATTTAGTCCAGCGCGCAT 720
Db |||||
221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY |||||
721 GCCTCTGAGCTTGTGATCCCAAGTACGAGCGCTACACATATCGTAACCAAGGCTGGCGCTCC 780
Db |||||
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY |||||
781 GTCGAGACCTCTGGGTGCTGAGGAGGAGTACCTCTGGTCTGTGTTATGCTTTGCATA 840
Db |||||
261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY |||||
841 CATGGCTCACTCGTAAATTCCTATATACTATACACCTATACCGGTGCGCTCGGGCTGTG 900
Db |||||
281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY |||||
901 GACTTTCGCTTGCAGCTTGCAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTTC 960
Db |||||
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY |||||
961 TCCGCTTATTCAGACACTGCTCGCACCGCTTGGTGGGTGCGGAGCGGAGTCCGAG 1020
Db |||||
321 SerArgTyrSerSerThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY |||||
1021 CTCACACACCGCTGCTACCGCTTTTGAAGGACCTCTATTTTACTAGTACTAATGTT 1080
Db |||||
341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY |||||
1081 GTCGCTGAGATCGCGCGGAGTACGCTTACCTGTTCAACCTTGTGACACTCTGCTT 1140
Db |||||
361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY |||||
1141 GCGGCGCTCGCACAGATGATTTGCTGCGGTGGCGAGCTGTTCTACTCCGCTCCC 1200
Db |||||
381 GlyGlyLeuProThrGluLeuIleSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY |||||
1201 GTTGTCTCAGCAATCGCGAGCGGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db |||||
401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY |||||
1261 CAGGATAGGGTATGCAATCCCGATGACATGACCTCGGAGATCTGCTGTGTTATT 1320
Db |||||
421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle 440
QY |||||
1321 CAGGATATGATATCAACATGAACAGATCGCGAGCGCTTCTCCAGCCCATCGCG 1380
Db |||||
441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY |||||
1381 CCTTTCTGCTTGCAGAGTAAATGATGCTTTGGCTCTCTCTCACCGCTCCGAGTAT 1440
Db |||||
461 ProPheSerValLeuArgAlaAsnAspValLeuTyrPheLeuSerLeuThrAlaAlaGluTyr 480
QY |||||
1441 GACCACTGCTATGGCTCTCGAGTGGCGCGCATGTTATGTTCTGACTCTGTGACCTTG 1500

481 AspGlnSerThrTyrGlySerThrGlyProValTyrValSerAspSerValThrLeu 500
QY |||||
1501 GTTATGTTGACCGCGCGCGAGCGCTTGGCGCGGAGTATGAGTGGACCAAGTGCACA 1560
Db |||||
501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY |||||
1561 CTTGACGGTCCGCTCTCCACCATCCAGCAGTACTCAAGACCTTCTTTGTCCTGCCG 1620
Db |||||
521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY |||||
1621 CTCGCGGTAGCTCTCTTTCTGGAGGCGAGCACAACTAAAGCCGGGTACCTTTATAAT 1680
Db |||||
541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY |||||
1681 TATAACACCACTGCTAGCGCAACTGCTTGTGAGAAATGCCCGCGGACCGGGTCCGT 1740
Db |||||
561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY |||||
1741 ATTTCACCTTACACCACTAGCTGCTGCTGCTCCGCTCTCCATTTCTCGGTTGCCGTT 1800
Db |||||
581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaAlaVal 600
QY |||||
1801 TTAGCCCCCACTCTGCGCTAGCATGCTTGAGGATACCTTGAGTACCTTCCGCGCGC 1860
Db |||||
601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY |||||
1861 CATACTTTTGATGATTTCTGCCAGAGTCCCGCGCTTGGCTTGGCTTCCAGGCTCGCTTC 1920
Db |||||
621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY |||||
1921 CAGTCTACTGCTCGCTGAGCTTCCAGCGCTTAAAGATGAAGGTGGGTAAAACTCGGGAGTTG 1980
Db |||||
641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 9
US-09-553-427-8
; Sequence 8, Application US/09553427
; Patent No. 6379891
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fy, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6379891-A/No. 6379891-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/553,427
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,507
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/336,672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0183.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-553-427-8

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-851-410a-6_COPY_5147_7129 (1-1983) x US-09-553-427-8 (1-660)

QY 1 ATGCGCCCTCGCCGCTATTTGTTGCTCTCCTCATGTTTTCCTGATGTCGCCGGCGCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCCGGCTCAGCCCTCGCCGGCTGTCGGCGCGCAGCGGGTTCGGCGGTGT 120
DB 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlySerGlyGly 40
QY 121 TTCTGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTATCCAAACCAAC 180
DB 41 PheTyrGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTGGCCCGATGTACCCGCTGGCGGGGCTGGACCTCGTGTTCGCAACCGGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTGGCTGACAGGCCCGCCGCGCCGCTTCCTCAGTCGT 300
DB 81 ArgProLeuGlySerAlaTyrArgAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACCTTACCACAGCTGGGCGCGCGCTAAACCGCGCTCGCTCCGCGCCCATGACACCCCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisPThrPro 120
QY 361 CCAGTGCCTGATGTGACCTCGCGCGCGCATCTTGGCGCGCGCAGTATTAACCTATCAACA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTCCGTCGGCCACCGGCACTAACTGCTTCTTTATGCGCCCTCTT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTTACCCCTTACGAGCGGCACCAATACCATATTAATGCGCACGGAAGCTTCT 540

DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTGCCGTCGCCCAATCCGTTACCGCCCGCTGTCGCCCAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCCGGCGTTACGCCATCTCCATCTCATCTCTGGCCACAGACCACCAACCCCGCAG 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATTAACCTCGACGAGTGTCTGATTTAGTCCAGCCCGCAT 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGTGGCGCTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrPargSer 260
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACTCTGGTCTTGTATTGCTTGCATA 840
DB 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTCGTAAATTCCTATATACTAATACCTATACCGGTGCGGCTCGGGCTGTG 900
DB 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGGCCTTGAGCTTGAGTTTCGCAACCTTACCGCCGCTAACCAACCAATACGGGTC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTTCAGCACTGCTCGCCACCGCTTCCTCGCGGTGCGGACGGGACTGCGGAG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACAGCGCTGCTACCGCTTTATGAGAGGACCTCTATTTACTAGTACTAATGTT 1080
DB 341 LeuThrThrThrAlaAlaThrArgPheMetIysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 CTCGCTGAGATCGCGCGGGATAGCCCTCACCTCTGTTCAACCTTGCCTGACACTCTGCTT 1140
DB 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCCACAGAAATTTGATTTCTGCGGTGTCGCGAGCTGTTCTACTCCCGTCCC 1200
DB 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
DB 401 ValValSerAlaAsnGlyGluProThrValIysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGTATTGCAATCCCGCATGACATTCGAGAGATCTCGTGTGGTTATT 1320
DB 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
QY 1321 CAGGATTATGATAACAACATGACAAGATCGCGCGAGCCCTTCTCCAGCCCATCGCC 1380
DB 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGCTTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT 1440
DB 461 ProPheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACGAGTCACTTATGGCTCTCGACTGCGCCAGTTTATGTTCTGACTCTGACCTGACCTTG 1500
DB 481 AspGlnSerThrTyrGlySerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGACCGCGCGCAGCGCTGTCGCGCTCGCTCGATTGGACCAAGGTCA 1560
DB 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrThrLysValThr 520
QY 1561 CTTGACGCTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTGCTCGCG 1620
DB 521 LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuPro 540

QY 1321 CAGGATTATGATAACCAACATGAACGATCGCGCGCTTCTCCAGCCCATCGCG 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProHisProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGTCTCTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCCACTATGGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTGACTTG 1500
Db 481 AspGlnSerThrTyrGlySerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGCGACCGCGCGAGCGCTTGGCCGCTCGCTGATTGGACCAAGGTCACA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTTGACGGTCGCCCTCTCCACCACTCCAGCAGTACTCGAAGACCTTTCTGTCTGCGCG 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGGCGGTAAAGTCTCTTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGCTAGCGACCACTGCTTGTGAGAATGCGCGCGGACCGGTCGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTCCACTTACCACTAGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTGCGCTAGCATGCTTGGAGGATACCTTGAGTACCTGCGCGCGCGCG 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACCTTTGATGATTCTGCCAGAGTGCCTGCGCGCGCTTGGCTTCCAGGCTTCTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGCTGAGCTTCCAGCGCTTAAGATGAAGTGGTAAACTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 11

US-07-870-985A-19
; Sequence 19, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Kwaczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07870,985A
; FILING DATE: 20-APRIL-1992

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0093.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
US-07-870-985A-19

Alignment Scores:

Pred. No.:	0	Length:	660
Score:	660.00	Matches:	660
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-07-870-985A-19 (1-660)

QY	1	ATGCGCCCTCGGCTATTTTGTGTGCTCTCTCATGTTTTCCTATGCTGCCGCGCA	60
Db	1	MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20
QY	61	CGCGCGGTACCGCTGTCGCGCGCTGTGGCGCGCGAGCGCGGTTCGCGCGGTGT	120
Db	21	ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlyGlyGly	40
QY	121	TTCTCGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCA	180
Db	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThr	60
QY	181	CCCTTCGCGCCCGATGTTCACCGCTCGCGCGGCTGGACCTCTGTTTCGCAACCCG	240
Db	61	ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnPro	80
QY	241	CGACCACTCGGCTCGCTTGGCGTGACCGAGCGCGCGCGCGCGTTCGCTTCACGT	300
Db	81	ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArg	100
QY	301	AGACCTACACAGCTGGCGCGCGCGGTAAACCGGGTCTCCGCGCGCATCACACCG	360
Db	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThr	120
QY	361	CCAGTGCCTGTGTCGACTCCCGCGCGCGCATCTTTCGCGCGCGCTATTAACCTAT	420

Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGTCGCCACCGCCTACCTGGTCTTATATGCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGCTTTTACCCCTTCAGACGGCCACCAATACCATATATATGCGCCACGGAAGCTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGTTGCCGTCGCAATCCGTTACCGCCCTGCTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGGGGTTAGCCATCTCCATCTCATCTCTGCGCCACACACACCGCCGAGC 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATAACCTCGACGGATGTTCTGTTATTTAGTCCAGCCCGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGATCCCAAGTAGCGCCCTACACTATCGTAACCAAGGTCGGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260
QY 781 CTCGAGACCTCTGGGGTGGCTCAGAGGAGGCTACTCTGGTCTGTTATGCTTGTGCATA 840
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCCTCCTGTAATTCCTATATAATACACCTATACCGGTGCCCTCGGGCTGTTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTGGCTTGGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCTCTATTCAGACACTCTCGCCACCGCTTCCTCGGGTGGCGGAGGAGTCCCGAG 1020
Db 321 SerArgTyrSerSerThrAlaAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACACACCGCTGCTACCCGCTTTATGAGGACCTCTATTTACTAGTACTAATGCT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCCGTGAGATCGCGCGGATAGCCCTCACCTTCACCTGTCGACACTCTGCTT 1140
Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCCGACAGATTGATTCGTCGGCTGGTGGCAGCTGTTCTACTCCCGTCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATGGCGAGCGACTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGTAATTGCAATCCGATGACATGACCTCGGAGATCTGCTGGTTATT 1320
Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
QY 1321 CAGGATTATGATAACCAACATCAAGATCGGCGAGCCCTTCTCCAGCCCATCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGCTTCGAGCTAATGATGCTGTTGCTCTCTCTCACCGCTCCGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACGAGTCCACTATGCTCTCGACTCGCCGCTTATGTTTCTGACTCTGCTGACCTTG 1500
Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500

QY 1501 GTTAATGTTGCGACCGCGCGCAGCGCTTGCCTCGCTCGTTCGATTGACCAAGTTCACA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspThrThrLysValThr 520
QY 1561 CTTGACGCTGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCCTGCGC 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAGCTCTCTTCTGGAGGAGCAGGACCACTAAAGCCGGGTACCTTATAT 1680
Db 541 LeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCATCTAGTAGCACAACCTGTTGTCGAGATGCGCGCGGACCGGGTCTGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACCATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAsnAlaAlaGlyHisArgValAla 620
QY 1861 CATACTTGTGATGATTTCTGCGCAGAGTCCCGCCCTTGGCTTTCAGGGCTGCGCTTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 12

PCT-US95-13703-13
; Sequence 13, Application PC/TUS9513703
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
; TITLE OF INVENTION: USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13703
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)

INDIVIDUAL ISOLATE: ORF-2	
PCT-US95-13703-13	
Alignment Scores:	
Pred. No.:	0
Score:	660.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	5
US-09-851-410A-6_COPY_5147_7129 (1-1983) x PCT-US95-13703-13 (1-660)	
QY	1 ATCGCGCCCTCGGCTATTTTGTGCTGCTCTCATGCTTTTGTGCTGCGCGCGCA 60
DB	1 MetArgProAArgProIleuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY	61 CGCCCGGTCAGCGTGTGCGCCCGCTGCGGCGCGCGCGCGGTTCCGCGGTGGT 120
DB	21 ProProGlyInProSerGlyArgArgArgGlyArgArgSerGlyGlyGly 40
QY	121 TTCTGGGTGACGGGTGATTTCTAGCCCTTCGCAATCCCTATATTATCAACCAAC 180
DB	41 PheTriPlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY	181 CCCTTCGCCCCGATGTCACCGCTGCGCGCGCGGCTGAGCCCTGCTGTGCGCAACCGGC 240
DB	61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY	241 CGACCACTCGGCTCCGCTTGGCGTGACCGCCAGCGCCCGCGCGCTTGCTCAGCTCGT 300
DB	81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY	301 AGACCTACACAGCTGGCGCGCGCTAAACCGCGGTGCTCCGCGCCATGACACCCCG 360
DB	101 ArgProThrThralaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY	361 CCAGTGCCTTGATCGACTCCGCGCGCGCATCTTCGCGCGGCGAGTAACTATCAACA 420
DB	121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
QY	421 TCTCCCTTTACCTCTTCGCTGGCCACCGCACTAACCTGCTTCTTATGCGCGCCCTCTT 480
DB	141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY	481 AGTCCGCTTTTACCTTACGAGCGGACCAATACCCATATAATGCGCACGGAAGTTCT 540
DB	161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY	541 AATTATGCCAGTACCGGGTGGCGTGCACCAATCCGTTACCGCCGCTGGTCCCAAT 600
DB	181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY	601 GCTGTCGGCGGTTACGCCATCTCCATCTCATTTCTGGCCACAGACCACCCCGGACG 660
DB	201 AlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrProThr 220
QY	661 TCGGTTCATAGNATTAACCTCGACGGATGTTGCTATTTAGTTCAGCCCGCGCAT 720
DB	221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY	721 GCCTCTGAGCTTGTGATCCCAAGTAGCGCTACACTATCTGTAACCAAGGCTGGCGCTC 780
DB	241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260
QY	781 GTCGAGACCTCTGGGTGGCTGAGGAGGCGCTACCTCTGGTCTGTGTTATGCTTTGCATA 840
DB	261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuGlyIle 280
QY	841 CATGGCTCACTCGTAATTCCTATACCTATACACCTTATACCGGTGCGCTCGGGCTGGT 900
DB	281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300

RESULT 13

QY	901 GACTTTGCGCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATACGCGGTC 960
DB	301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY	961 TCCGTTATTCCAGCACTCTCGCCACCGCTTCTGTCGCGGTGCGAGCGGACTGCGGAG 1020
DB	321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY	1021 CTCACCAACACGGCTCTACCCGCTTTATGTAAGGACCTCTATTATTACTAGTACTATGGT 1080
DB	341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY	1081 GTCGCTGACATCGCGCGCGGATAGCCCTCACCTGTTCAACCTTCTGACACTCTGCTT 1140
DB	361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY	1141 GCGCGCTCCGACAGAAATGATTTCTGTCGCTGGTGGCGAGCTGTTCTACTCCCGTCCC 1200
DB	381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY	1201 GTTGCTCAGCAATGCGGAGCGGCTGTAAGTTGTATACATCTCTAGAGATGCTCAG 1260
DB	401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY	1261 CAGGATAAGGCTATTGCAATCCGCAATGACATTGACCTCGAGAAATCTCGTGTGGTTATT 1320
DB	421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle 440
QY	1321 CAGGATTATGATAACCAACATGAACATCGCGCGAGCTTCTCCAGCCCAATCGCGC 1380
DB	441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY	1381 CCTTTCTCTCTCGAGCTAATGATGCTGTTGCTCTCTCTCAGCTGCGGAGTAT 1440
DB	461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY	1441 GACCACTCCACTTATGGCTCTTCGACTGGCGCGAGTTTATGTTTCTGACTCTGTGACTTG 1500
DB	481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY	1501 GTTAATGTTTCGACCGCGCGCGCGCTGTCGCTCGATTGGACCAAGGTCA 1560
DB	501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY	1561 CTTGAGGTCGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTCTTTGTCCTGCGC 1620
DB	521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY	1621 CTCGCGGGTAAAGCTCTTTTCTGGGAGGCGAGCACAACTAAAGCCGGGTACCCCTTAAAT 1680
DB	541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY	1681 TATAACACCACTGCTAGCAACCAACTGCTTGTGCGAATGCGCGCGCGGACCGGGTCGCT 1740
DB	561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY	1741 ATTTCCACTTACCACTAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1800
DB	581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY	1801 TTAGCCCCCACTCTGCGCTAGCATTGCTTGGAGTACCTTGGACTACCTGCGCCCGCGC 1860
DB	601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAlaAspThrLeuAspTyrProAlaArgAla 620
QY	1861 CATACTTTTGTGATTTCTGCCAGAGTCCGCCCTTTGGCTTTCAGGGCTTCGCGTTTC 1920
DB	621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY	1921 CAGTCTACTGTCGTCAGCTTCAGCGCTTACAGTGAAGCTGGCTAAACTCGGAGTTG 1980
DB	641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

US-08-542-634-15
; Sequence 15, Application US/08542634
; Patent No. 6214970
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; APPLICANT: McAtee, C. Patrick
; APPLICANT: Yarbough, Patrice O.
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/542,634
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 549 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)
; INDIVIDUAL ISOLATE: r62kDa, FIGURE 4
US-08-542-634-15

Alignment Scores:
Pred. No.: 0 Length: 549
Score: 549.00 Matches: 549
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.18% Indels: 0
DB: 3 Gaps: 0

US-09-851-410a-6_COPY_5147_7129 (1-1983) x US-08-542-634-15 (1-549)
QY 334 GCGTCCGCGGAGTAAACCTATCAACATCTCCCTTACCTCCGTCGGCCACCGGCAC 393
Db 1 AlaValAlaProAlaHisAspThrProProValProAspValAspSerArgGlyAla 20
QY 394 TTGCGCGCGGAGTAAACCTATCAACATCTCCCTTACCTCCGTCGGCCACCGGCAC 453
Db 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThr 40
QY 454 AACCTGGTCTTTATGGCGGCCCTTAGTCCGCTTTTACCCCTTACGAGCGGACCAAT 513
Db 41 AsnLeuValLeuTyrAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACCCATATATGGCCACGGAAGCTTCTAAATTATGCCAGTACCGGGTTGCCCGGCCACA 573
Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThr 80
QY 574 ATCCGTTACCGCGCGCTGGTCCCCAAATGCTGTGCGCGGTACGCCATCTCCATCTCATTC 633

81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100
QY 634 TGGCCACAGACACACCCAGCGTCGCTGGTATATGAATCAATAACCTCAGCGAT 693
Db 101 TrpProGlnThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp 120
QY 694 GTTCGTATTTTGTCCAGCCCGCATAGCTCTGAGCTTGTGATCCCAAGTAGCGGCTA 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGlnLeuValIleProSerGluArgueu 140
QY 754 CACTATCGTTAAACCAAGGCTCGCTCGAGACCTCTCGGGTGGTGGAGGAGGCT 813
Db 141 HisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCTGGTCTTGTATGCTTTCATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerLeuValAsnSerTyrThrAsnThr 180
QY 874 CCCTATACCGGTGCGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200
QY 934 ACCCGCGGTAAACACCAATACCGCGGCTCTCCGTTTATTCAGCACTGCTGCCACGCTT 993
Db 201 ThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
QY 994 CGTCGCGGTGCGGACCGGACTGCGGAGCTCACACACGCGCTGCTGCTGCTGCTGCTGCTGCT 1053
Db 221 ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrThrAlaAlaThrArgPheMetLys 240
QY 1054 GACCTCTATTTTACTAGTACTAATGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1113
Db 241 AspLeuTyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThr 260
QY 1114 CTGTTCAACCTTGTGACACTCTGCTTGGCGGCTGCCGACAGAAATTTGATTTCTCGCT 1173
Db 261 LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAla 280
QY 1174 GGTGGCCAGCTGTTCTACTCCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1233
Db 281 GlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLys 300
QY 1234 TTGTATACATCTGTAGAGATGCTCAGCAGATAGGGTATTGCAATCCCGCATGACATT 1293
Db 301 LeuTyrThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspIle 320
QY 1294 GACCTCGGAGAACTCTGTTGTTTATTCAGGATTATGATAACCAACATGAACAAGATCGG 1353
Db 321 AspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArg 340
QY 1354 CCGACGCTTCTCCAGCGCCATCGCGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1413
Db 341 ProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeu 360
QY 1414 TGGCTCTCTCTCACCGTCCGAGTATGACAGTCCACTTATGCTCTTCCAGTGGGCCA 1473
Db 361 TrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPro 380
QY 1474 GTTTATGTTTCTGACTCTGTGACTTGTGCTTAATGTTGCGACCGCGCGGCGGCTGGCC 1533
Db 381 ValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla 400
QY 1534 CGGTGCTCCATGGACCAAGGTACACTTTCAGGTGCGGCTCCCTCTCCACCATCCACGAG 1593
Db 401 ArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln 420
QY 1594 TACTCGAAGACCTTCTTTTGTCTCCGCTCCGCGGTAAAGTCTCTTCTTCTGGAGGCGAGC 1653
Db 421 TyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGly 440
QY 1654 ACAACTAAAGCGGGTACCTTATTAATTAACCACTGCTAGCGACCACTGCTGCTTC 1713

Db 441 ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuVal 460
QY 1714 GAGATGCGCGCGGACCGGTCGTATTTCCACTTACACCTAGCCTGGTGTGT 1773
Db 461 GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGly 480
QY 1774 CCGCTCTCCATTCTCGCGTTCGCGTTTGTAGCCGCCCACTCTCGCTAGCATTTGCTTGAG 1833
Db 481 ProValSerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuLeuGlu 500
QY 1834 GATACCTTGGACTACCTGCGCGGCCCATCTTTTGATGATTTCTGCCAGAGTCCGC 1893
Db 501 AspThrLeuAspTyrProAlaArgAlaHisThrPheAspPheCysProGluCysArg 520
QY 1894 CCGCTTGGCTTCAGGCTCGGCTTTCAGTCTACTGTGCTGAGCTTCAGCGCCTTAAG 1953
Db 521 ProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLys 540
QY 1954 ATGAGGTGGTAAACTCGGAGGTG 1980
Db 541 MetLysValGlyLysThrArgGluLeu 549

RESULT 14

US-08-477-292-15
; Sequence 15, Application US/08477292
; Patent No. 6291641
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas
; APPLICANT: McAttee, Patrick
; APPLICANT: Yarbough, Patrice
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cathleen M. Desjardins, M.D.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: US
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,292
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/327,952
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, Allan A
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: G32P5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)369-9500
; TELEFAX: (415)368-0709
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 549 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)
; INDIVIDUAL ISOLATE: r62kDa, FIGURE 4
US-08-477-292-15
Alignment Scores:

Pred. No.: 0 Length: 549
Score: 549.00 Matches: 549
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.18% Indels: 0
DB: 3 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-08-477-292-15 (1-549)
QY 334 GGGTGGTCCGGCCCATGACACCCCGCAGTGCCTGATGTGACTCCCGCGCGCATC 393
Db 1 AlaValAlaProAlaHisAspThrProValProAspValAspSerArgGlyAlaIle 20
QY 394 TTGCGCGGCACTATAACCTATCAACATCTCCCTTACCTCTTCGTGGCCACCGCACT 453
Db 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThr 40
QY 454 AACCTGGTTCTTTATGCGCCCGCTCTTATGTCGCTTTTACCCCTTCAGACGCGCAAT 513
Db 41 AsnLeuValLeuTyrAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACCATATAATGCGCACGGAAGCTTCTAATTATGCCAGTACGGGTTCGCGTGCACA 573
Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThr 80
QY 574 ATCCGTACCGCCCGCTGGTCCCAATGCTGCGGGTTAGCCCATCTCCATCTCATTC 633
Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyTyrAlaIleSerIleSerPhe 100
QY 634 TGGCCACAGACACCAACCCCGAGCTCCGTTGATATGAATTCATAACCTCGACGAT 693
Db 101 TrpProGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120
QY 694 GTTGTATTTTATGTCAGCCCGGATAGCTCTGAGCTTGTGATCCCAAGTAGCGCCTA 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTATCGTAACCAAGGTGCGCTCCGTCGAGACCTCTGGGTGCTCAGGAGGAGGCT 813
Db 141 HisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCTGGTCTTTGTTATGCTTTGCATACATGGCTCCTCGTAAATTCCTATATAACA 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerLeuValAsnSerTyrThrAsnThr 180
QY 874 CCGTATACCGGTGCCCTCGGCTGTGTGACTTTGGCTTGGCTTGGCTTGGCTTGGCTT 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200
QY 934 ACCCGGTAAACCAATACGCGGTCTCCGTTATTCAGACACTGCTCGCACCGCCTT 993
Db 201 ThrProGlyAsnThrAsnThrArgValSerArgTyrSerThrThrAlaArgHisArgLeu 220
QY 994 CGTGGGTGCGGACGGGACTGCGGAGCTCACACCGGCTGCTACCGCTTTTATGAAG 1053
Db 221 ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLys 240
QY 1054 GACCTCTATTTACTAGTACTAATGTTGCTGAGATCGGCGCGGATAGCCCTCACC 1113
Db 241 AspLeuTyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThr 260
QY 1114 CTGTTCAACCTTGTGACACTCTGCTTGGCGGCTCGCCAGACAGAAATTGATTCGTCGCT 1173
Db 261 LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuLeuSerAla 280
QY 1174 GGTGGCCAGCTGTTCTACTCCCGTCCGCTGTGCTCAGCCAATGGCGAGCGGCTGTAAG 1233
Db 281 GlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLys 300
QY 1234 TTGTATACATCTGTAGAGAAATGCTCAGCAGGATAGGGTATTGCAATCCCGATGACATT 1293
Db 301 LeuTyrThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIle 320

```
QY 1294 GACCTCGAGAAATCTCTGTGGTTATTACAGGATTATGATACCAACCAATGAACAGATCGG 1353
Db 321 AspleuGlyGlnSerArgValValleGlnAspTyrAspGlnHisGluGlnAspArg 340
QY 1354 CCGACGGCTTCTCCAGCCCATCGCGCCCTTCTCTGTCTCCAGCTAATGATGCTT 1413
Db 341 ProthrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeu 360
QY 1414 TGGCTCTCTCACCCTGCGGAGTAGACAGTCCACTTATGCTCTTCCAGTGGCCCA 1473
Db 361 TrpleuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPro 380
QY 1474 GTTTATGTTCTGACTCTGTGACCTTGTGTAATGTTGCGCGCGCGAGCGGTTGCC 1533
Db 381 ValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla 400
QY 1534 CGGTCTCTCGATGGACCAAGGTGCACACTTTGACGGTGCAGCGGCTCTCCACCATCCAGCAG 1593
Db 401 ArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln 420
QY 1594 TACTCGAAGACCTCTTTGCTGCTGCGCTCGCGGTAAAGCTCTCTTCTCGGAGGCGGC 1653
Db 421 TyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGly 440
QY 1654 ACACTAAAGCCGGGTACCCCTTATAATTATAACACCACTGCTAGCGACCAACTGCTTGTG 1713
Db 441 ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuVal 460
QY 1714 GAGAAATCGCGCGGACCGGGTCCGCTATTTCCTACCTTACCACTACGCTGGGTGGT 1773
Db 461 GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGly 480
QY 1774 CCGCTCTCCATTCTCGGGTTCGGTTCCTAGCCCGCCACTGCGCTAGCATTCCTTGAG 1833
Db 481 ProValSerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGlu 500
QY 1834 GATACCTTGACCTACCCCTGCGCGCCCATCTTTGATGATTTCTGCCAGAGTGCCTG 1893
Db 501 AspThrLeuAspTyrProAlaArgAlaHisThrPheAspAspPheCysProGluCysArg 520
QY 1894 CCCCTTGCGCTTCAGGCTGCGCTTCAGTCTACTGCTGCTAGCTTACGCGCTTAAG 1953
Db 521 ProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLys 540
QY 1954 ATGAAGTGGGTAAACTCGGGAGTTG 1980
Db 541 MetLysValGlyLysThrArgGluLeu 549
```

RESULT 15

PCT-US95-13703-15
; Sequence 15, Application PC/TUS9513703

GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
TITLE OF INVENTION: USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13703
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.41
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 549 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)
INDIVIDUAL ISOLATE: r62kDa, FIGURE 4
PCT-US95-13703-15
Alignment Scores:
Pred. No.: 0 Length: 549
Score: 549.00 Matches: 549
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.18% Indels: 0
DB: 5 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x PCT-US95-13703-15 (1-549)

```
QY 334 GCGTCTCGCTCCGGCCCATGACACCCCGCAGTGCCTGATGCGACTCCCGGGCGCATC 393
Db 1 AlavaAlaProAlaHisAspThrProValProAspValAspSerArgGlyAlaIle 20
QY 394 TTGGCCGGCAGATATAACCTATCAACATCTCCCTTACCTCTCCGTGGCCACCGGCACT 453
Db 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThr 40
QY 454 AACCTGTTCTTTATGCGCGCCCTCTTATGTCGGTTTTTACCCCTTACAGGACGGCACCAAT 513
Db 41 AsnLeuValLeuTyrAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACCCATATAATGCCACGGAAGCTTCTTAATTATGCCAGTACCGGGTTCGCCGTGCCACA 573
Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThr 80
QY 574 ATCGGTTACCGCGCTGGTCCCCCATGCTGTGCGCGGTTACGCCATCTCCATCTCATTC 633
Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100
QY 634 TGGCCACAGACACACCCACCGACGTCGTTGATATGAATTCATAAATCTCGACCGAT 693
Db 101 TrpProGlnThrThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp 120
QY 694 GTTCGTATTTTACTCCAGCCCGCATAGCTCTGAGCTTGTGATCCCAAGTAGCGCCTA 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTATCGTAACCAAGCTCGCTCCGTGAGACCTCTGGGTGGCTGAGGAGGAGCT 813
Db 141 HisTyrArgAsnGlnGlyTrpArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCTGGTCTTGTATGCTTTGCATACATGGCTCACTCGTAAATTCCTATCTAATACA 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerLeuValAsnSerTyrThrAsnThr 180
QY 874 CCCTATACCGGTCCTCGGGCTGTGGACTTTGCCCTTCCAGCTTGTAGTTTCGCACTT 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200
QY 934 ACCCCGGTAACACCAATACCGGGTCTCCGTTATTCAGCACTGCTCCACCGGCTT 993
Db 201 ThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
```

QY 994 CGTCGGGTGGAGCGGAGCTCCGAGCTCACCACCGGTGCTACCGCTTTATGAAG 1053
Db |||||||
QY 221 ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLys 240
Db |||||||
QY 1054 GACCTCTATTTTACTAGTAAATGGTTCGGTGAATCGGCGGAGATAGCCCTCACC 1113
Db |||||||
QY 241 AspLeuTyrPheThrSerThrAsnGlyValGlyGluLeuGlyArgGlyIleAlaLeuThr 260
Db |||||||
QY 1114 CTGTTCAACCTTGTGACACTCTGCTGGCGGCTCCGACAGAAATGATTCGTCGGCT 1173
Db |||||||
QY 261 LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAla 280
Db |||||||
QY 1174 GGTGCCAGCTGTTCTACTCCGTCCTCCGTTGCTCAGCAATGGCGAGCGGACTGTTAAG 1233
Db |||||||
QY 281 GlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLys 300
Db |||||||
QY 1234 TTGTATACATCTGTAGAGAATGCTCAGCAGGATAAGGGTATTGCAATCCGCGATCACATT 1293
Db |||||||
QY 301 LeuTyrThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIle 320
Db |||||||
QY 1294 GACCTCGAGAAATCTCGTGTGTTATTTCAGGATTATGATAACCAACATGAACAAGATCGG 1353
Db |||||||
QY 321 AspLeuGlyGluSerArgValIleGlnAspTyrAspAsnGlnHisGlnGlnAspArg 340
Db |||||||
QY 1354 CCGAGCGCTTCTCCAGCCCATCGCGCCCTTCTCTGTCCTTCGAGCTAATGATGTGCTT 1413
Db |||||||
QY 341 ProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeu 360
Db |||||||
QY 1414 TGGCTCTCTCTCACCGTCCGAGTATGACCAAGTCCACTTATGGTCTTTCGACTGGCCCA 1473
Db |||||||
QY 361 TrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerThrGlyPro 380
Db |||||||
QY 1474 GTTATGTTCTTGACTCTGTGACCTTGGTATTATCTGGACCGGCGGAGCGGCTTGGC 1533
Db |||||||
QY 381 ValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla 400
Db |||||||
QY 1534 CGGTCCGCTCGATTGACCAAGGTCACTTGACGGTCCGCCCTCTCCACCATCCAGCAG 1593
Db |||||||
QY 401 ArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln 420
Db |||||||
QY 1594 TACTCGAAGACCTCTCTTGTCTCGGCTCCGCGTAAAGTCTCTCTTCTGGAGGAGGC 1653
Db |||||||
QY 1654 ACACTAAAGCGGTACCTTATATATATACACCACTGCTAGCACCACTGCTTGTTC 1713
Db |||||||
QY 441 ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuVal 460
Db |||||||
QY 1714 GAGATGCCCGCGGCGGCTCGCTATTTCACCTTACACCACTAGCCTGGGTGCTGGT 1773
Db |||||||
QY 461 GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGly 480
Db |||||||
QY 1774 CCCGTCCTCCATTCTCGGTTCCGTTTTCGCTTTCAGCCCCCACTCTCGCTAGCATGCTTGAG 1833
Db |||||||
QY 481 ProValSerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGlu 500
Db |||||||
QY 1834 GATACCTTGACACTACCTCCCGGCGGCTACTTTTGATGATTTCTGCCAGAGTCCCGC 1893
Db |||||||
QY 501 AspThrLeuAspTyrProAlaArgAlaHisThrPheAspPheCysProGluCysArg 520
Db |||||||
QY 1894 CCCCTTGGGCTTCAGGGCTCGCTTTCACAGTCTACTGCTGCTGAGCTTCAGGCGCTTAAG 1953
Db |||||||
QY 521 ProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLys 540
Db |||||||
QY 1954 ATGAGGTGGTAAACTCGGGAGTTG 1980
Db |||||||
QY 541 MetLysValGlyLysThrArgGluLeu 549
Db |||||||

Search completed: August 16, 2004, 13:59:39
Job time : 57.0151 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:55:54 ; Search time 97.3236 Seconds
(without alignments)

12792.733 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129

Perfect score: 660

Sequence: 1 ATGCGCCCTCGGCTATT.....GTAAACTCGGAGTTGTAG 1983

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1292805 seqs, 313927144 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2482574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlsp
-Q=/cgn2_1/USPTO_spool_p/US09851410/runat_16082004_125719_9771/app_query.fasta_1.4238
-DB=Published Applications AA -QPMT=fastan -SUFFIX=oli.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40 cdi -LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09851410@cgn_1_1_27@runat_16082004_125719_9771
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
---------------	-------	----------------	--------	----	-------------

1	660	100.0	660	9	US-09-769-066-13	Sequence 13, Appl
2	660	100.0	660	10	US-09-851-410-8	Sequence 8, Appl
3	660	100.0	660	14	US-10-185-868-19	Sequence 19, Appl
4	549	83.2	549	9	US-09-769-066-15	Sequence 15, Appl
5	541	82.0	660	15	US-10-257-044-1	Sequence 1, Appl
6	540	81.8	540	9	US-09-769-066-25	Sequence 25, Appl
7	525	79.5	525	9	US-09-769-066-27	Sequence 27, Appl
8	436	66.1	436	14	US-10-165-868-17	Sequence 17, Appl
9	408	61.8	660	12	US-10-381-770-1	Sequence 1, Appl
10	327	49.5	327	9	US-09-769-066-17	Sequence 17, Appl
11	327	49.5	327	14	US-10-165-868-15	Sequence 15, Appl
12	210	31.8	210	12	US-10-381-770-2	Sequence 2, Appl
13	210	31.8	232	12	US-10-381-770-12	Sequence 12, Appl
14	200	30.3	389	12	US-10-381-770-4	Sequence 4, Appl
15	190	28.8	194	12	US-10-381-770-3	Sequence 3, Appl
16	165	25.0	169	12	US-10-381-770-7	Sequence 7, Appl
17	155	23.5	159	12	US-10-381-770-8	Sequence 8, Appl
18	145	22.0	149	12	US-10-381-770-9	Sequence 9, Appl
19	135	20.5	139	12	US-10-381-770-10	Sequence 10, Appl
20	120	18.2	525	9	US-09-769-066-28	Sequence 28, Appl
21	120	18.2	540	9	US-09-769-066-26	Sequence 26, Appl
22	120	18.2	549	9	US-09-769-066-16	Sequence 16, Appl
23	120	18.2	659	14	US-10-165-868-20	Sequence 20, Appl
24	120	18.2	660	9	US-09-769-066-14	Sequence 14, Appl
25	118	17.9	674	16	US-10-239-090A-55	Sequence 55, Appl
26	118	17.9	674	16	US-10-239-090A-57	Sequence 57, Appl
27	109	16.5	123	10	US-09-851-410-9	Sequence 9, Appl
28	109	16.5	123	12	US-10-381-770-11	Sequence 11, Appl
29	109	16.5	124	9	US-09-769-066-21	Sequence 21, Appl
30	78	11.8	123	14	US-10-224-999A-3478	Sequence 3478, Ap
31	75	11.4	660	10	US-09-468-147-92	Sequence 92, Appl
32	75	11.4	660	12	US-10-319-745-92	Sequence 92, Appl
33	75	11.4	660	16	US-10-239-090A-51	Sequence 51, Appl
34	75	11.4	660	16	US-10-239-090A-53	Sequence 53, Appl
35	61	9.2	124	9	US-09-769-066-22	Sequence 22, Appl
36	61	9.2	327	10	US-09-468-147-175	Sequence 175, App
37	61	9.2	327	10	US-09-468-147-176	Sequence 176, App
38	61	9.2	327	12	US-10-319-745-175	Sequence 175, App
39	61	9.2	327	12	US-10-319-745-176	Sequence 176, App
40	61	9.2	338	10	US-09-468-147-199	Sequence 199, App
41	61	9.2	338	12	US-09-468-147-200	Sequence 200, App
42	61	9.2	338	12	US-10-319-745-199	Sequence 199, App
43	61	9.2	338	12	US-10-319-745-200	Sequence 200, App
44	61	9.2	459	10	US-09-468-147-206	Sequence 206, App
45	61	9.2	459	10	US-09-468-147-207	Sequence 207, App

ALIGNMENTS

RESULT 1
US-09-769-066-13
; Sequence 13, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; Mcatee, C. Patrick
; Yarbough, Patrice O.
; Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

```
; APPLICATION NUMBER: US/09/769,066
; FILING DATE: 24-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/542,634
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; STRANDEDNESS: Hepatitis E Virus (Burma strain)
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
;
; US-09-769-066-13
;
; Alignment Scores:
; Pred. No.: 0 Length: 660
; Score: 660.00 Matches: 660
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 100.00% Indels: 0
; DB: 9 Gaps: 0
;
; US-09-851-410a-6_COPY_5147_7129 (1-1983) x US-09-769-066-13 (1-660)
;
; QY 1 ATCGCGCTCGGCTATTTTGTCTGCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
; DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
;
; QY 61 CCGCGCGCTCAGCGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGTTCGGGGTGGT 120
; DB 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly 40
;
; QY 121 TTTCTGGGTGACCGGTTGATTTCTCAGCCTTCGCAATCCCTATATTCATCAACCAAC 180
; DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProIleHisProThrAsn 60
;
; QY 181 CCCTTCGCGCGCGCGCTCAGCGCTGCGCGCGCGCGCGCTGCGCTGCTGCGCGCGCGCC 240
; DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
;
; QY 241 CGACGACTCGGCTCCGCTTGGCTGACCGCGCGCGCGCGCGCGCGCGCGCTGCTGCTGCT 300
; DB 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
;
; QY 301 AGACCTTACCACAGCTGCGCGCGCGCGCGCTAACCGCGCTGCTGCGCGCGCGCGCGCG 360
; DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
;
; QY 361 CAGTGCCTGATGTCAGCTCCGCGCGCGCGCATCTTGGCGCGCGCGAGTAACTATCAACA 420
; DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTrpAsnLeuSerThr 140
;
; QY 421 TCTCCCTTACCTTCCGCGCGCGCGCGCGCTAACCTGCTTCTTATGCGCGCGCGCTTCT 480
; DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTrpAlaAlaProLeu 160
;
; QY 481 AGTCCGCTTTTACCTTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCT 540
; DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
;
; QY 541 AATTATGCCAGTACCGGGTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAAT 600
;
; Db 181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
;
; QY 601 GCTGTGGCGGTTACGCCATCTCCATCTCTATCTGGCCACAGACCACACCCCGGACG 660
;
; Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
;
; QY 661 TCCGTTGATATGAATTCATAACCTCGACGAGTGTTCGTATTTTAGTCCAGCCGCGCAT 720
;
; Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
;
; QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGCTCGCGCTCC 780
;
; Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
;
; QY 781 GTCGAGACCTCTGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
;
; Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
;
; QY 841 CATGGCTCACTCGTAAATTCCTATATACATACACCTATACCGGTGCGCTCGGCTGTTG 900
;
; Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
;
; QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACTTACCCCGGTAAACCAATACCGGCTC 960
;
; Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
;
; QY 961 TCCGCTATTTCAGCACTGCTCGCCACCGCTTCGTCGGTGGCGGCGGAGCTCGCGAG 1020
;
; Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
;
; QY 1021 CTCACCAACACGCGCTGCTACCGCTTTATGAAGAGCTCTATTTTACTAGTACTAATGGT 1080
;
; Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
;
; QY 1081 GTCGCTGAGATCGGCGCGGATAGCCCTCACCTGTTTCACTGCTGCTGCTGCTGCTGCT 1140
;
; Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
;
; QY 1141 GCGCGCTCGCGACAGATTGATTCGTCGGCTGGTGGCGGCTGCTTCTACTCCGCTCCC 1200
;
; Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
;
; QY 1201 GTTGTCTCAGCAATCGGCGCGGCTGTTAACTGTATATCTGTATACATCTGTAGAGATGCTCAG 1260
;
; Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
;
; QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAACTCTGTGGTTATT 1320
;
; Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle 440
;
; QY 1321 CAGGATTATGATAACCAACATGAACAGATCGCGCGCGCTTCTCCAGCGCCCATCGCGC 1380
;
; Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
;
; QY 1381 CCTTTCTCTCTCTCGAGCTAATGATGTGCTTTGGCTCTCTCTCTCAGCGCTCGCGAGTAT 1440
;
; Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyr 480
;
; QY 1441 GACGAGTCACTTATGGCTCTTCGACTGGCGCGAGTTATGTTTCTGACTCTGACCTTG 1500
;
; Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
;
; QY 1501 GTTAATGTTTCGACCGCGCGCGGCTGCGCGGCTCGCTCGGCTCGGCTCGGCTCGGCT 1560
;
; Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
;
; QY 1561 CTTGACGCTGCGCGCTCTCCACCAATCCAGCACTCTCGAAGACCTTCTTTGCTCTCGCG 1620
;
; Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
;
; QY 1621 CTCGCGGTAAAGCTCTTCTCGGAGGAGGAGGAGCACTAAAGCGGCTACCTTATAAT 1680
```

Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyProTyAsn 560
QY 1681 TATACACACCTAGGACCAACTGCTTGTGCGAATGCCCGGCGACCGGTCGCT 1740
Db 561 TyAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaGlyHisArgValAla 580
QY 1741 ATTTCACACTACACACTAGCCTGGGTGCTGCTCCGCTCCATTCTCGGTTGCCGTT 1800
Db 581 IleSerThrTyThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCCTAGCGCTAGCATGCTTGGATGAGTACCTGCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyProAlaArgAla 620
QY 1861 CATACCTTTGTGATTTCTGCCAGAGTGCCGCCCTTGGCTTCAGGGCTGCGCTTC 1920
Db 621 HisThrPheAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGAGCTTCAGCGCCTTAAGATGAAGTGGTAAACTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 2
US-09-851-410-8
; Sequence 8, Application US/09851410
; Publication No. US20030124510A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; Yarbrough, Patrice O
; Bradley, Daniel W
; Krawczynski, Krzysztof Z
; Tam, Albert
; Fy, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,410
; FILING DATE: 07-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/128,275
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: US 07/505,898
; FILING DATE: 05-APR-1990
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-851-410-8
Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-851-410-8 (1-660)
QY 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTTCCTATGCTGCGCGCCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCCCGGTTCAGCCGCTCTGGCGCGCTGTCGGCGCGCGCGCGGTTCGGCGGTGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly 40
QY 121 TTCTGGGTGACCGGGTGAATCTCAGCCCTTCGCAATCCCTATATTCATCAACCA 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyIleHisProThr 60
QY 181 CCCTTCGCGCCCGATGTCACCGCTCGCGCGGGGTGACCTCGTGTTCGCAACCCG 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGln 80
QY 241 CGACCACTCGGCTCGCTTGGCGTGACCAAGGCCCGCGCGCGCTTCACGTGCT 300
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSer 100
QY 301 AGACTTACACAGCTGGGGCGCGCGCTAACCGGGTGTCTCGCGCGCGCATGACAC 360
Db 101 ArgProThrThrAlaAlaProLeuThrAlaValAlaProAlaHisAspThr 120
QY 361 CCAGTGCCTGATGTCGACTCCGCGCGCGCCATCTTGGCGCGCGCATATAACCTA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnIleAsnLeuSer 140
QY 421 TCTCCCTTACTCTCTTCCGTGCGCACCGGCACTAACCTGGTTCTTTTATGCGCC 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyAlaAla 160
QY 481 AGTCGCTTTTACCCCTTCAGGCGGCGCAATACCATATAATGCGCGCAAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGlu 180
QY 541 AATATGCCAGTACCGGGTTCGCGTCCCAATCCGTTACCGCGCGCTGTGCTCCCA 600
Db 181 AsnTyAlaGlnTyArgValAlaArgAlaThrIleArgTyArgProLeuValPro 200
QY 601 GCTGTGCGCGGTACGCCATCTTCATCTCATCTCTGGCCACAGACACCAACCCG 660
Db 201 AlaValGlyTyAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCAATAACCTTCAGCGGATGTTCTGATTTAGTCAGCCG 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnPro 240
QY 721 GCCTCTGAGCTTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGT 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyArgAsnGlnGlyTrpArg 260
QY 781 GTCGAGACCTCTGGGGTGGCTCAGGAGGAGGCTACCTGCTGTTGTTATGCTTTG 840

Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
 QY 841 CATGGCTCACTCTAATTCCTATATACACCTATACCGGTGCCCTCGGGCTGTG 900
 Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
 QY 901 GACTTTGGCCCTCAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTC 960
 Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
 QY 961 TCCCGTTATTCCAGCACTGCTCCGCCACCGCTTCGTCGGGTCCGGACGGAGTCCGAG 1020
 Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
 QY 1021 CTACACACACGGCTGCTACCCGCTTTATGAAGACCTCTATTTTACTAGTACTATGGT 1080
 Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
 QY 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTTGTCTGACACTCTGCTT 1140
 Db 361 ValGlyGluLeuGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
 QY 1141 GCGCGCTCCGACAGAAATTGATTCGTGGCTGGTGGCAGCTGTCTACTCCCGTCC 1200
 Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyLeuLeuPheTyrSerArgPro 400
 QY 1201 GTTGTCTCAGCAATCGCGACGACTGTGTTAAGTTGTATATCATCTGTAGAGATGCTCAG 1260
 Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
 QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAACTCTCGTGGTTATT 1320
 Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
 QY 1321 CAGGATTATGATAACCAACATGAACAGATCGCGACGCTTCTCCAGCCCATCGCGC 1380
 Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
 QY 1381 CTTTCTCTCTCTCGAGTATGATGCTGTTGGTCTCTCTCACCGGTGCGGAGTAT 1440
 Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
 QY 1441 GACCACTGCTTATGCTCTCGACTGGCCCACTTATGCTTCTGACTCTGTGACCTTG 1500
 Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
 QY 1501 GTTAATGTCGACCGCGCGAGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1560
 Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
 QY 1561 CTTGACGGTCCGCTCTCCACATCCAGCACTGCTGAGACCTCTTTGCTGCTGCGG 1620
 Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
 QY 1621 CTCGCGGTGAAGCTCTCTTTCTGGGAGGACCACTAAAGCCGGTACCTCTTATAAT 1680
 Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
 QY 1681 TATAACACACTGCTAGCGACCACTGCTTGTGAGAAATCGCGCGGACCGCGGTGCT 1740
 Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
 QY 1741 ATTTCCACTTACACCTAGCTGGGTGCTGGTCCGCTCTCCATTTCTGGGTGCGGTT 1800
 Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
 QY 1801 TTAGCCCCCACTCTGCGTAGCATGCTTGGAGATACCTTGGACTACCTCGCGCGGCC 1860
 Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
 QY 1861 CATACTTTGATGATTTCTGCCAGAGTGGCGGCTTGGCTTCCGCTTACGGCTGCGCTTC 1920
 Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640

QY 1921 CAGTCTACTGTCTCAGCTTCAGCGCCTTAAGATGAAGTGGGTAAAACTCGGAGTTG 1980
 Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 3
 US-10-165-868-19
 ; Sequence 19, Application US/10165868
 ; Publication No. US20030143241A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reyes, Gregory R.
 ; Bradley, Daniel W.
 ; Twu, Jr-Shin
 ; Purdy, Michael A.
 ; Tam, Albert W.
 ; Krawczynski, Krzysztof Z.
 ; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Dehlinger & Associates
 ; STREET: 350 Cambridge Avenue, Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/165,868
 ; FILING DATE: 06-Jun-2002
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/870,985A
 ; FILING DATE: 20-APRIL-1992
 ; APPLICATION NUMBER: US 822,335
 ; FILING DATE: 17-JAN-1992
 ; APPLICATION NUMBER: US 505,888
 ; FILING DATE: 05-APRIL-1990
 ; APPLICATION NUMBER: US 420,921
 ; FILING DATE: 13-OCTOBER-1989
 ; APPLICATION NUMBER: US 367,486
 ; FILING DATE: 16-JUNE-1989
 ; APPLICATION NUMBER: US 336,672
 ; FILING DATE: 11-APRIL-1989
 ; APPLICATION NUMBER: US 208,997
 ; FILING DATE: 17-JUNE-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fabian, Gary R.
 ; REGISTRATION NUMBER: 33,875
 ; REFERENCE/DOCKET NUMBER: 4600-0093.30
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 324-0880
 ; TELEFAX: (415) 324-0960
 ; INFORMATION FOR SEQ ID NO: 19:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 660 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
 ; US-10-165-868-19

Alignment Scores:
 Pred No.: 0 Length: 660
 Score: 660.00 Matches: 660
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-165-868-19 (1-660)			
QY	1	ATGGCCCTCGCCGCTATTTGTTGCTGCTCTCATGTTTTCCTATGCTGCGCGGCCA	60
DB	1	MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20
QY	61	CGGCCGCTACGCGTCTGCGCGCGCTGTTGGCGCGGAGCGGGTTCGGCGGTGT	120
DB	21	ProProGlyGlnProSerGlyArgArgArgGlyArgSerGlyGlySerGlyGly	40
QY	121	TTCTGGGGTGACCGGTTGATTCTCAGCCCTTCGAATCCCTATATTCATCCACCAAC	180
DB	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn	60
QY	181	CCCTTCGCCCGCCGATGTCCACCGCTCGCGCGGGCTGGACCTCTGTTTCGCCAACCCGCC	240
DB	61	ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla	80
QY	241	CGACACTCGGCTCGCTTGGGTGACACAGGCCCGAGCGCGCGCTGCTGCTCAGTGT	300
DB	81	ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArg	100
QY	301	AGACCTACACAGCTGGCGCGCGCTTAACCGCGCTCGCTCGCGCCCATGACACCCCG	360
DB	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThr	120
QY	361	CCAGTCCCTGATGTGCACTCCCGCGCGCCATCTTTGCGCGCGCAGTATACCACTA	420
DB	121	ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr	140
QY	421	TCTCCCTTACTCTTCGTTGCGCCACCGGCACTAACCTGGTCTTATGCGCGCCCTCT	480
DB	141	SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaPro	160
QY	481	AGTCGGCTTTACCCCTTCAGACGCGCACCATACCATATATATGCGCACGGAAGCTCT	540
DB	161	SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAla	180
QY	541	AATTATGCCAGTACCGGGTTCGCCGTGCCAATCCGTTACCGCGCGTGGTCCCAAT	600
DB	181	AsnTyrAlaGlnTyrArgValAlaAlaAlaThrIleArgTyrArgProLeuValPro	200
QY	601	GCTGTCGGGGTTACGCATCTCCATCTCTGCGCCACAGACACACACCGCGAGC	660
DB	201	AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr	220
QY	661	TCCGTTGATATCAATTCATACCTCGAGGATGTTCTGTTATTTAGTCCAGCCCGCAT	720
DB	221	SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGly	240
QY	721	GCTCTGAGCTGTGATCCCAAGTAGCGCCPACACTATCGTAACCAAGCTGGCGCTCC	780
DB	241	AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArg	260
QY	781	GTCGAGACTCTGGGGTGGCTGAGAGGAGGCTACTCTGCTGCTTGTATGCTTTCGCA	840
DB	261	ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle	280
QY	841	CATGCTCACTCGTAATTCCTATACCTACACCTATACCGGTGCCCTCGGCTGTG	900
DB	281	HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeu	300
QY	901	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTC	960
DB	301	AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArg	320
QY	961	TCCGTTATCCAGCACTGCTGCCACCGCTTGGTGGCGGTGCGGACCGGAGCTGCCG	1020
DB	321	SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAla	340
QY	1021	CTCACCAACGCGGTGCTACCCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGT	1080
DB	341	LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsn	360
QY	1081	GTCGGTAGATCGCGCGGGATAGCCCTCACCCCTGTTCAACCTTGCTGACACTCTGCT	1140
DB	361	ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeu	380
QY	1141	GGCGCCTGCGCAGACAGAAATTCATTTCGTCGCTGGTGGCCAGCTCTTACTCCCGTCC	1200
DB	381	GlyGlyLeuProThrGluLeuIleSerAlaGlyGlnLeuPheTyrSerArgPro	400
QY	1201	GTGTCTCAGCAATGGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAACTCTC	1260
DB	401	ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAla	420
QY	1261	CAGATAAGGTATTGCAATCCCGCATGCAATTGACCTCGGAGAACTCTGTTGTTATT	1320
DB	421	GlnAspLysGlyIleAlaIleProHisAspLeuGlyGluSerArgValIle	440
QY	1321	CAGGATATGATAACCAACATGAAACATCGCGCGCCCTCTCCAGCCCATCGCGC	1380
DB	441	GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSer	460
QY	1381	CCTTTCTCTGCTCCTTCGAGCTAATGATGCTGCTTTGGCTCTCTCTCACCGCTGCC	1440
DB	461	ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGlu	480
QY	1441	GACAGTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTG	1500
DB	481	AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThr	500
QY	1501	GTTAATGTTGCGCGCGCGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1560
DB	501	ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysVal	520
QY	1561	CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTCTCTTCTCCTCG	1620
DB	521	LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuPro	540
QY	1621	CTCCGCGTAGCTCTCTTTCTGGAGGAGGAGCAACAATAAGCGGGTACCTTATATAT	1680
DB	541	LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyr	560
QY	1681	TATAACACACTCTGAGCGACCACTGCTGTCGAGAAATGCGCGCGGACCGGTCGCT	1740
DB	561	TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgVal	580
QY	1741	ATTTCCACTTACACCACTAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1800
DB	581	IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAla	600
QY	1801	TTAGCCCCCACTCTGCGCTAGCATGCTTGAGGATACCTTGGACTACCTGCCCGCGC	1860
DB	601	LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArg	620
QY	1861	CATACCTTTGATGATTTCTGCGCAGTGGCGCGCCCTTGGCTTTCAGGCTCGCTTC	1920
DB	621	HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAla	640
QY	1921	CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAAGTGAAGGTGGTAAACTCGGAGTTG	1980
DB	641	GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu	660
RESULT 4			
US-09-769-066-15			
; Sequence 15, Application US/09769066			
; Patent No. US20020107360A1			
; GENERAL INFORMATION:			
; APPLICANT: Fuerst, Thomas R.			
; McAtee, C. Patrick			
; Yarbrough, Patrice O.			
; Zhang, Yifan			

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/769,066
FILING DATE: 24-Jan-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/542,634
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.30

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:

LENGTH: 549 amino acids
TYPE: amino acid

STRANDEDNESS: Hepatitis E virus (Burma strain)
r62kDa, FIGURE 4

TOPOLOGY: linear
MOLECULE TYPE: protein

HYPOTHETICAL: NO
ORIGINAL SOURCE:

SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-769-066-15

Alignment Scores:

Pred. No.: 0 Length: 549
Score: 549.00 Matches: 549
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.18% Indels: 0
DB: 9 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-769-066-15 (1-549)

QY	334	CGGGTCTCCGGCCCATGACACCCGCGAGTGCCTGATGTCGACTCCGCGCGGCATC	393
Db	1	AlaValAlaProAlaHisAspThrProValProAspValAspSerArgGlyAlaIle	20
QY	394	TTGCGCGCGGAGTAACTATCAATCTCCCTTACCTTCCGCTGGCCACCGGCACCT	453
Db	21	LeuArgArgGlnIyrAsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThr	40
QY	454	AACCTGGTCTTTATGCGCGCCCTCTTAGTCGCTTTTACCCTTCAGGACGCGCAAT	513
Db	41	AsnLeuValLeuIyrAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsn	60
QY	514	ACCCATATATGCCACGAGNAGTCTTAATTATGCCAGTACCGGTTGCCGTGCCACA	573
Db	61	ThrHisIleMetAlaThrGluAlaSerAsnIyrAlaGlnIyrArgValAlaArgAlaThr	80
QY	574	ATCCGTATACCGCGCTGTGCCCAATGCTGTGCGCGGTTACGCCATCTCCATCTCATTC	633
Db	81	IleArgIyrArgProLeuValProAsnAlaValGlyGlyIyrAlaIleSerIleSerPhe	100
QY	634	TGGCCACAGACACACCCGCGAGTCCGCTGGATATGAATCAATCACTCGAGGAT	693

Db	101	TrpProGlnThrThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp	120
QY	694	GTTCGTATTATTAGTCCAGCCGCGCATAGCTCTGAGCTTGTGATCCCAAGTAGCGCTA	753
Db	121	ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu	140
QY	754	CACATATCGTAACCAAGGCTGGGCTCCGTCGAGACCTCTGGGGTGGCTGAGGAGGCT	813
Db	141	HisIyrArgAsnGlnGlyIyrArgSerValGluThrSerGlyValAlaGluGluAla	160
QY	814	ACCTCTGGTCTTGTATGCTTTGCATACATGGCTCACTCGTAAATTCCTATATAACA	873
Db	161	ThrSerGlyLeuValMetLeuCysIleHisGlySerLeuValAsnSerIyrThrAsnThr	180
QY	874	CCCTATACCGTGCCTCGGCTGTGGACTTTGGCTTGCCTTGAGCTTGAGTTCCCACTT	933
Db	181	ProIyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuPheArgAsnLeu	200
QY	934	ACCCCGGTAAACACCAATACCGGGTCTCCCGTTATTCCAGCACTGCTGCCACCGCTT	993
Db	201	ThrProGlyAsnThrAsnThrArgValSerArgIyrSerSerThrAlaArgHisArgLeu	220
QY	994	CGTCCGGTCCGAGCGGACTCCGAGCTCACCAACCGCTCTACCCGCTTTATGAAG	1053
Db	221	ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLys	240
QY	1054	GACCTCTATTACTAGTACTAATGCTGCTGGTGAGATCGGCGCGGGATAGCCCTCACC	1113
Db	241	AspLeuIyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThr	260
QY	1114	CTGTTCACACTCTGCTGACACTCTGCTGGCGGCTCCGACAGAAATGATTTCTCGTGGCT	1173
Db	261	LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAla	280
QY	1174	GGTGGCCAGCTGTTCTACTCCCGTCCCGTGTCTCAGCAATCGCAGCGAGCTGTTAAG	1233
Db	281	GlyGlyGlnLeuPheIyrSerArgProValValSerAlaAsnGlyGluProThrValLys	300
QY	1234	TTGTATACATCTGTAGAGATGCTCAGCAGGATAGGATTTCGAATCCCGCATGACATT	1293
Db	301	LeuIyrThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIle	320
QY	1294	GACCTCGGAGATCTCGTGTGTTATTTCAGGATTATGATAACCAACATGAACAGATCGG	1353
Db	321	AspLeuGlyGluSerArgValValIleGlnAspIyrAspAsnGlnHisGluGlnAspArg	340
QY	1354	CCGACGCTTCTCCACGCCCATCGCGCCCTTTCTGTCTCTCGAGCTAATGATGCTTT	1413
Db	341	ProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeu	360
QY	1414	TGGCTCTCTCTCACCGCTCCGAGTATGACAGTCCACTTATGGCTCTTCGACTGGCCCA	1473
Db	361	TrpLeuSerLeuThrAlaAlaGluIyrAspGlnSerIyrGlySerSerThrGlyPro	380
QY	1474	GTATTATGTTCTGACTCTCTGACCTTGGTTAATGTTGCCACCGCGCGGCGGCTGCC	1533
Db	381	ValIyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla	400
QY	1534	CGGTCGCTCGATTGGACCAAGGTACACTTGACGGTCCCGCTCTCCACCATCCAGCAG	1593
Db	401	ArgSerLeuAspIyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln	420
QY	1594	TACTCGAAGACCTCTTTGCTCGCGCTCCGCGTAAGCTCTCTTCTGGGAGGCGAGC	1653
Db	421	TyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheIyrGluAlaGly	440
QY	1654	ACAACATAACCGGGTACCTTATATATAACACCACTTCGCTAGCGCAACACTGCTTGC	1713
Db	441	ThrThrLysAlaGlyIyrProIyrAsnIyrAsnThrThrAlaSerAspGlnLeuLeuVal	460
QY	1714	GAGAATCCCGCGGCGGCGGCTCGCTATTTCCACTTACCACTAGCTGGGTGGTGGT	1773
Db	461	GluAsnAlaAlaGlyHisArgValAlaIleSerThrIyrThrThrSerLeuGlyAlaGly	480

```
QY 1774 CCGGCTCCATTTCTGCGGTTGCGGTTTGTAGCCCGCCACTCTGCGCTAGCATTTGCTGAG 1833
DB 481 ProValSerIleSerAlaValAlaValIleuAlaProHisSerAlaLeuAlaLeuLeuGlu 500
QY 1834 GATACCTTGGACTACCTCCGCGCCGCCATACCTTTTGATGATTTCTGCGCAGAGTCCGCGC 1893
DB 501 AspThrLeuAspTyrProAlaAlaAlaHisThrPheAspAspPheCysProGluCysArg 520
QY 1894 CCCCTTGGCTTCAGGGCTGCGCTTCCAGTCTACTGCTGCTGAGCTTCAGCGCCTTAAG 1953
DB 521 ProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLys 540
QY 1954 ATGAAGGTGGTAAACTCGGGAGTTG 1980
DB 541 MetLysValGlyLysThrArgGluLeu 549

RESULT 5
US-10-257-044-1
; Sequence 1, Application US/10257044
; Publication No. US20030220475A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; APPLICANT: Represented by the Secretary, Department of Health and
; APPLICANT: Human Services, Centers for Disease Control and Prevention
; APPLICANT: HOWARD A. FIELDS
; APPLICANT: YURI E. KHUDYAKOV
; APPLICANT: JIHONG MENG
; TITLE OF INVENTION: NEUTRALIZING IMMUNOGENIC HEV POLYPEPTIDES (HEV)
; FILE REFERENCE: 14114.034002
; CURRENT APPLICATION NUMBER: US/10/257,044
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,380
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 660
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: / No. US20030220475A1e =
US-10-257-044-1
Alignment Scores:
Pred. No.: 0 Length: 660
Score: 541.00 Matches: 641
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 81.97% Indels: 0
DB: 15 Gaps: 0

US-09-851-410a-6_copy_5147_7129 (1-1983) x US-10-257-044-1 (1-660)
QY 55 GCGCACCGCGGFCAGCGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGGTTCCGCGC 114
DB 19 AlaProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGly 38
QY 115 GGTGTTTCTGGGTGACCGGTTGATCTCAGCCCTTCGCAATCCCTTATATTCATCCA 174
DB 39 GlyGlyPheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisPro 58
QY 175 ACCAACCCCTTCGCCCCGATGTCACGCTGCGCGCGCGGCTGAGCCTGTTTCGCCAA 234
DB 59 ThrAsnProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGln 78
QY 235 CCGCGCCGACCACTCGGCTCGCTTGGGTGACGAGGCCCGCGCGCGCTGCTGCTCA 294
DB 79 ProAlaArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAsp 98
QY 295 CGTCGTAGACCTACCACTGGGCGCGCGCTAACCGCGCTGCTCGCGCGCTGCTGAC 354
```

```
DB 99 ArgArgArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAsp 118
QY 355 ACCCGCCGAGTGCCTGATGTCGACTCCGCGCGCGCATCTTCGCGCGGAGATTAACCTA 414
DB 119 ThrProProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeu 138
QY 415 TCAACATCTCCCTTACCTCTTCCGTGGCCACCGGCACCTAACTGCTGTTCTTTATCGCGC 474
DB 139 SerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAla 158
QY 475 CTTCTTACTGCGCTTTTACCCCTTCAGACGCGCACCAATACCATATATATGTCACGCGAA 534
DB 159 ProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGlu 178
QY 535 CTTTCTAATTATGCCCCAGTACCGGTTGCGCGTGGCCACAATCCGTTACCGCGCGTGTGC 594
DB 179 AlaSerAsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuVal 198
QY 595 CCCAATGCTGTCGCGGTTTACGCCATCTCCATCTCATTTCTGGCCACAGACCACCCACC 654
DB 199 ProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThr 218
QY 655 CGACGCTCGGTTGATGATGAATTCATAACCTCGACGAGTGTTCGTAATTTAGTCACGCGC 714
DB 219 ProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnPro 238
QY 715 GGCATAGCTCTGAGCTTGTGATCCCAAGTAGCGCCCTACACTATCGTAACCAAGGCTGG 774
DB 239 GlyIleAlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrp 258
QY 775 CGTCCGTCGAGACCTCTGGGTCGCTGAGGAGGAGGTACCTCTGCTGCTTCTTATGCTT 834
DB 259 ArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeu 278
QY 835 TGCATACATGGCTCACTCGTAAATTCCTATACATAACACCCCTATACCGGTCCTCGGG 894
DB 279 CysIleHisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGly 298
QY 895 CTGTTGGACTTTGCCCTTGAGCTTGAGTTGCAACCTTACCCCGGTAACCAACCAATAGC 954
DB 299 LeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThr 318
QY 955 CGGTCCTCCGCTTATTCGACGACACTCTCGCCACCGCTTCTGTCGCGGTGCGACGCGACT 1014
DB 319 ArgValSerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThr 338
QY 1015 GCGAGCTCACACACCGCTGCTACCGCTTATGAAGGACCTCTATTTTACTAGTACT 1074
DB 339 AlaGluLeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThr 358
QY 1075 AATGTCGTCGTCGATCGCGCGGGATAGCCCTCACCCCTGTTCAACCTTGTCTGACACT 1134
DB 359 AsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThr 378
QY 1135 CTGCTTGGCGGCTCCCGACAGAAATTTGATTCGTCGCTGGTGGCGAGCTGTTCTACTCC 1194
DB 379 LeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyCysGlnLeuPheTyrSer 398
QY 1195 CGTCCCGTTGTCTCAGCCAAATGGCGAGCCGACTGTTAAAGTTGTATACATCTGTAGAGAT 1254
DB 399 ArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsn 418
QY 1255 GCTCAGGAGGATAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGT 1314
DB 419 AlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgVal 438
QY 1315 GTTATTCAGGATATGATAACCAACATGAACAAGATCGCGCGAGCTTCTCCAGCCCA 1374
DB 439 ValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaPro 458
QY 1375 TCGCGCCCTTTCTCTGTCTCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCTCACCGCTGCC 1434
DB 459 SerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAla 478
```



```
QY 694 GTTCGTAATTTAGTCCAGCCCGCATAGCTCTGAGCTTGATCCCAAGTGAGCGCTA 753
Db 121 ValArgIleuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTATCGTAACAACAGCGCTGGCGCTCGTCGAGACCTCTGGGGTGGCTGAGGAGGAGCT 813
Db 141 HisTyrArgAsnGlnGlyTrpArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACTCTGGTCTGTGTTATGCTTTCATACATAGGCTCACTCGTAAATTCCTATACATAACA 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerLeuValAsnSerTyrThrAsnThr 180
QY 874 CCTATACCGTCCCTCGCGCTGTTGGACTTTCCTTGCCTTGCCTTGCCTTGCCTTGCCTT 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200
QY 934 ACCCCCGGTAAACACCAATACGCGGTCTCCCGTTATTCAGACACTCTCGCCACCGCCTT 993
Db 201 ThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
QY 994 CGTCGGGTGCGGACGGGACTCCGAGCTCACCAACCGGCTGCTACCGCTTATGAAG 1053
Db 221 ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLys 240
QY 1054 GACCTCTATTTACTAGTACTAATGTGTGGTGAGATCGCCGCGGATAGCCCTCAC 1113
Db 241 AspLeuTyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThr 260
QY 1114 CTGTTCAACCTTGCTGACACTCTGCTTGGCGGCTCGCCAGACAGATTGATTCGTCGGCT 1173
Db 261 LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAla 280
QY 1174 GGTGGCCAGCTGTTCTACTCCCGTCCCGTTCCTCAGCAATGCGCAGCGAGCTTAAG 1233
Db 281 GlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLys 300
QY 1234 TTGTATACATCTGTAGAGATGCTCAGCAGGATAGGGTATTGCAATCCGGATGACATT 1293
Db 301 LeuTyrThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIle 320
QY 1294 GACCTGGAGATCTGCTGTGTTATTCAGGATTAATGATAACCAACATGAACAGATCG 1353
Db 321 AspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArg 340
QY 1354 CGAGCGCTTCCAGCCCGCTGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1413
Db 341 ProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeu 360
QY 1414 TGGCTCTCTCTCACCGCTCGCGATGATGACCAATGCTTATGGCTCTCTCTCTCTCTCT 1473
Db 361 TrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPro 380
QY 1474 GTTTATGTTCTGACTCTGTGACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1533
Db 381 ValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla 400
QY 1534 CGGTGCTCGATTGGACCAAGTGCACATTGACGGTCCCGCTCTCCACCATCCAGCAG 1593
Db 401 ArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln 420
QY 1594 TACTCGAAGACCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1653
Db 421 TyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGly 440
QY 1654 ACAACTAAAGCGGGTACCTTATATTAATATACACACTGCTAGGACCACTGCTTGTCT 1713
Db 441 ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuVal 460
QY 1714 GAGATGCGCGCGGACCGGGTCTGCTATTTCACATTACACACTAGCTGCTGCTGCTGCT 1773
Db 461 GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGly 480
```

1774 CCGCTCTCATTTCTCGGTTGCGGTTTGTAGCCCCCACTCTCGCTAGCATTTGCTTGG 1833
|||||
481 ProValSerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuGlu 500
|||||
1834 GATACCTTGAGCTACCTCGCGCGCCCATACTTTTGATGATTTCTGCCAGAGTCCCGC 1893
|||||
501 AspThrLeuAspTyrProAlaArgAlaHisThrPheAspAspPheCysProGluCysArg 520
|||||
1894 CCGCTTGGCTTCAG 1908
|||||
521 ProLeuGlyLeuGln 525
|||||

RESULT 8
US-10-165-868-17
; Sequence 17, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; Bradley, Daniel W.
; Twu, Jr-Shin
; Purdy, Michael A.
; Tam, Albert W.
; Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/165,868
; Filing Date: 06-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; Filing Date: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; Filing Date: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; Filing Date: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921
; Filing Date: 13-OCTOBER-1989
; APPLICATION NUMBER: US 367,486
; Filing Date: 16-JUNE-1989
; APPLICATION NUMBER: US 336,672
; Filing Date: 11-APRIL-1989
; APPLICATION NUMBER: US 208,997
; Filing Date: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 436 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 9
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:

US-10-165-868-17

Alignment Scores:

Pred. No.: 0 Length: 436
Score: 436.00 Matches: 436
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.06% Indels: 0
DB: 14 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-165-868-17 (1-436)

QY 673 AATTCAATAAAGCTGACGATGTTTCGATTTTAGTCCAGCCGCGCATAGCCCTCAGCTT 732
DB 1 AsnSerIleThrSerThraspValArgIleLeuValGlnProGlyIleAlaSerGluLeu 20
QY 733 GTGATCCCAAGTAGCGCTACACTATCGTAAACCAAGGCTGGCGCTCCGTGAGACCTCT 792
DB 21 ValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThrSer 40
QY 793 GGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGTCATACATGGCTCACTC 852
DB 41 GlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySerLeu 60
QY 853 GTAAATTCCTATACATAACACCTATACCGGTGCCCTCGGGCTGTTGGACTTTTCCCTT 912
DB 61 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 80
QY 913 GAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATACCGGGTCTCCGGTTATTCC 972
DB 81 GluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTyrSer 100
QY 973 AGCACTGCTCGCCACCGCTTCGTCCGGTGGGAGCGGAGTCCGAGCTACACACCAG 1032
DB 101 SerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGluLeuThrThr 120
QY 1033 GCTGTACCGCTTATCAAGGACCTCTATTTTACTAGTAAATGGTTCGGTCGAGATC 1092
DB 121 AlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGlyValGlyGluLe 140
QY 1093 GGCCCGGGATAGCCCTCACCCCTGTTCACCTTGCTGACACTCTCTGGCGGCTCGCG 1152
DB 141 GlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuPro 160
QY 1153 ACAGAAATGATTCGTGGCTGGTGGCCAGCTGTTCTACTCCGCTCCGCTTGTCTCAGC 1212
DB 161 ThrGluLeuIleSerSerAlaGlyGlnLeuPheTyrSerArgProValValSerAla 180
QY 1213 AATGGCAGCGGACGTTTAAAGTTGTATACATCTGTAGAGATGCTCAGCAGATAGGGT 1272
DB 181 AsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGlnAspLysGly 200
QY 1273 ATTGCAATCCCGCATGACATTGACCTCGGAGAACTCTCGTGTGTTATTTCAGGATTATGAT 1332
DB 201 IleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIleGlnAspTyrAsp 220
QY 1333 AACCAACATGACAGATCGCCGAGCTTCTCAGCCCGCATCGGCCCTTCTCTGTC 1392
DB 221 AsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerVal 240
QY 1393 CTTTCAGCTTAATGATGTCTTTGGCTCTCTCACCGCTGCCGAGTATGACAGTCCACT 1452
DB 241 LeuArgAlaAsnAspValLeuIlePheSerLeuThrAlaAlaGluTyrAspGlnSerThr 260
QY 1453 TATGGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTGGTTTAAATGTGG 1512
DB 261 TyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnValAla 280
QY 1513 ACCGGCGCGAGCGCTTCCCGCTGGCTCGATGAGCAAGGTCACACTTGCAGCGGTGCG 1572
DB 281 ThrGlyValAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArg 300
QY 1573 CCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTCCGCTCCGCGTAA 1632

DB 301 ProLeuSerThrIleGlnGlnTyrSerLysThrPheValLeuProLeuArgGlyLys 320
QY 1633 CTCTCTTTCTGGAGGAGGAGGACAACTAAAGCCGGTACCCTTATAATATTAACACCACT 1692
DB 321 LeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThr 340
QY 1693 GCTAGCGACCAACTGCTTGTGAGAATCCCGCGGACCGGGTGGCTATTTCCTACTTAC 1752
DB 341 AlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyr 360
QY 1753 ACCACTAGCCTGGGTGGTGGTCCCGCTCTCCATTTCTGCGGTGGCGTTTATAGCCCCCAC 1812
DB 361 ThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaValLeuAlaProHis 380
QY 1813 TCTGGCTAGCATTCCTTGAGATACCTTGACTACCTCCGCGGCGCCCATACTTTTGTAT 1872
DB 381 SerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAlaHisThrPheAsp 400
QY 1873 GATTTCCTCCAGAGTCCGCCCTTCGCTTCAGGGCTCGCTTTCAGTCTACTACTGTC 1932
DB 401 AspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrVal 420
QY 1933 GCTGAGCTTCAGCGCTTAAAGATGAAGTGGGTAAACTCGGGAGTTG 1980
DB 421 AlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 436

RESULT 9

US-10-381-770-1

; Sequence 1, Application US/10381770

; Publication No. US20040052813A1

; GENERAL INFORMATION:

; APPLICANT: YANG SHENG TANG company, Ltd.

; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPOS

; FILE REFERENCE: IEC010037PCT

; CURRENT APPLICATION NUMBER: US/10/381,770

; CURRENT FILING DATE: 2003-07-16

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 660

; TYPE: PRT

; ORGANISM: hepatitis E virus

US-10-381-770-1

Alignment Scores:

Pred. No.: 0 Length: 660
Score: 408.00 Matches: 508
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 61.82% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-381-770-1 (1-660)

QY 289 GCCTCAGCTCGTAGACTACACAGCTGGGCGCGCGCTAACCGCGTCCCTCGGCG 348
DB 97 AlaSerArgArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAla 116
QY 349 CATGACACCCCGCAGTGCCTGATGCTGACTCCCGCGCGCCATCTTGGCGCGGAGTAT 408
DB 117 HisAspThrProProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyr 136
QY 409 AACCTATCAACATCTCCCTTACCTCTTCGTTGGCCACCGCAGCTAACCTGGTCTTTAT 468
DB 137 AsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyr 156
QY 469 GCCGCGCCCTCTAGTCCGCTTTTACCCCTTACGAGGAGGACCAATACCATATATGCCC 528
DB 157 AlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAla 176
QY 529 ACGGAAGCTTCTTAATTATGATGCCAGTACCGGGTGGCGGTCGCCACAAATCGTTACCGCG 588

177 ThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgPro 196
589 CTGGTCCCAATGCTGTGGCGGTAGCGCATCTCCATCTCATCTGTGGCCACAGACACC 648
197 LeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThr 216
649 ACCACCCGACGTCGGTGTATGAATTCATTAACCTCGACGGATGTTCTGATTTAGTC 708
217 ThrThrProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuVal 236
709 CAGCCCGCATAGCTCTGAGCTTGTGATCCCAAGTCAGCGCTTACACTATCGTAACCAA 768
237 GlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGln 256
769 GGTGGCGCTCCGTCGAGACTCTGGGGTGGCTGAGGAGGAGGAGTCTGCTGTCTTGT 828
257 GlyTyrArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuVal 276
829 ATGCTTTGCATACATGCTCACTCGTAAATTCCTACTATAATACACCTATACCGGTGCC 888
277 MetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAla 296
889 CTCGGGCTGTGAGCTTTGCGCTTGAAGTTGAGTTTCGCAACCTTACCCCGGTAAACACC 948
297 LeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThr 316
949 AATACGGGGTCTCCCGTTATTCAGCACTGTCGCGACCGCTTGTGCGGTGGGAC 1008
317 AsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAsp 336
1009 GGGACTGCGGAGCTCACACACGCTGCTACCGCTTATGAAGGAGCTCTATTTTACT 1068
337 GlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetIysAspLeuTyrPheThr 356
1069 AGTACTAATGCTGCTGAGATCGCGCGGATGAGCTTCAACCTGTTTCAACCTTGCT 1128
357 SerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAla 376
1129 GAACTCTGTTGGCGCTGCGCACAGATTAATGCTGCGCTGTCGCGGAGCTGTTTC 1188
377 AspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPhe 396
1189 TACTCCGCTCCGCTGCTCAGCAATGGCGAGCGACTGTTAAGTTGTATACATCTGTA 1248
397 TyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyrThrSerVal 416
1249 GAGATGCTCAGCAGATTAAGGTATTGCAATCCCGCATGACATTCACCTCGGAGATCT 1308
417 GluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSer 436
1309 CGTGTGTTATTTCAGATTTATGATACCAACATGACATGACATTCACCTCGGAGATCT 1368
437 ArgValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPro 456
1369 GCCCATCGCGCTTCTCTGCTTCGAGCTAATGATGCTGTTGGCTCTCTCTCACC 1428
457 AlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuThrIleLeuSerLeuThr 476
1429 GTTGGCAGATGACCAAGTCCATTATGGCTTCGACTGCGCCAGTTATGTTTCTGAC 1488
477 AlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAsp 496
1489 TCTGTGACCTTGTTTAAATGTTGGCGCGCGCGCTGCGCGCTGCTGCTGCTGCTG 1548
497 SerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyr 516
1549 ACCAGGTCACTTACAGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTC 1608
517 ThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhe 536
1609 TTTGCTCTGCGCTCCGCGTAACTCTCTTCTGGAGGAGGAGGAGCAACTAAAGCGGG 1668

537 PheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGly 556
1669 TACCCTTATAATATTAACACCACTGCTAGGACCAACTGCTTCTGAGAAATGCGCGGG 1728
557 TyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGly 576
1729 CACCGGTCGCTATTTCCACTTACACCACTAGCCTGGGTGCTGCTGCTGCTGCTGCT 1788
577 HisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSer 596
1789 GCGTTGCCGCTTTAGCCCCCACTCT 1815
597 AlaValAlaValLeuAlaProHisSer 605
RESULT 10
US-09-769-066-17
; Sequence 17, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; McAttee, C. Patrick
; Yarbrough, Patrice O.
; Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09769,066
; FILING DATE: 24-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/542,634
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 amino acids
; TYPE: amino acid
; STRANDEDNESS: Hepatitis E Virus (Burma strain) SG3
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-769-066-17
Alignment Scores:
Pred. No.: 2,51e-293 Length: 327
Score: 327.00 Matches: 327
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.55% Indels: 0
Gaps: 0
DB:
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-769-066-17 (1-327)

QY 1000 GGTGGGACGGAGCTGCGAGCTCACCACCGGCTGCTACCGCTTTTATGAGGACCTC 1059
Db 1 GlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetIysAspLeu 20
QY 1060 TATTTTACTAGTACTAATGGTTCGGTGAGATCGCGCGGGATAGCCCTCACCTGTTC 1119
Db 21 TyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPhe 40
QY 1120 AACCTTGTGACACTCTGCTGGCGGCTGCGGACAGAAATTGATTCGTGCGCTGTGGC 1179
Db 41 AsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGly 60
QY 1180 CAGCTGTTCTACTCCGCTGCTGCTCAGCAATGGCGAGCGCACTGTTAAGTTGTAT 1239
Db 61 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 80
QY 1240 ACATCTGTAGAGATGCTCAGCAGGATAAGGTATTGCAATCCCGCATGACATTCACCTC 1299
Db 81 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 100
QY 1300 GGAGATCTCGTGGTATTATTCAGGATATGATACCAACATGACAAAGATCGCGGAGC 1359
Db 101 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 120
QY 1360 CCTTCTCCAGCCCCATCGCGCCCTTCTCTGCTCCTCGAGCTAATGATGCTGTTGGCTC 1419
Db 121 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrPhe 140
QY 1420 TCTCTCAGCGTCCGAGPATCACCAGTCCACTTATGGCTCTTCGACTGGCCAGTTTAT 1479
Db 141 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 160
QY 1480 GTTTCGTGACTCTGACCTTGTGTTAATGTTGGACCGCGCGCGAGCGCTTCCCGGTGC 1539
Db 161 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 180
QY 1540 CTCGATTGGACCAAGCTCACACTTCAGCGTCCCGCTCTCCACCATCCAGCAGTACTCG 1599
Db 181 LeuAspThrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 200
QY 1600 AAGACCTTTCTTGTCTGCTGCGGTCCGCGTAAAGCTCTCTTCTGGGAGGACGACAACT 1659
Db 201 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 220
QY 1660 AAAGCGGGTACCTTATATTAATTAACACACTGCTAGGACCACTGCTTTCAGAAAT 1719
Db 221 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 240
QY 1720 GCCGCGGGCACCGGTGCTATTTCCACTTACACCATGAGTCTGGGTGCTGCTCCCGTC 1779
Db 241 AlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal 260
QY 1780 TCCATTTCTCGGGTCCGCTTTAGCCCCCACTCGCGCTAGCATTCGCTCAGGATACC 1839
Db 261 SerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuGluAspThr 280
QY 1840 TTGGACTACCTGCGCGGCCCATACTTTTGATGATTTCTGCCAGAGTCCGCCCTT 1899
Db 281 LeuAspTyrProAlaArgAlaHisThrPheAspAspPheCysProGluCysArgProLeu 300
QY 1900 GGCCTTCAGGGTGGCTTTCAGTCTACTGTCGTGAGCTTCAGCGCTTAAAGATGAAG 1959
Db 301 GlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuIysMetIys 320
QY 1960 GTGGGTAAACTCGGAGTTG 1980
Db 321 ValGlyLysThrArgGluLeu 327

RESULT 11

US-10-165-868-15

; Sequence 15, Application US/10165868

; Publication No. US20030143241A1

GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R.
Bradley, Daniel W.
Tsu, Jr-Shin
Purdy, Michael A.
Tam, Albert W.
Krawczynski, Krzysztof Z.
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/165,868
FILING DATE: 06-Jun-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/870,985A
FILING DATE: 20-APRIL-1992
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0093.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 327 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: SG3, BURMA, FIGURE 9
SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-165-868-15

Alignment Scores:
Pred. No.: 2,51e-293 Length: 327
Score: 327.00 Matches: 327
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.55% Indels: 0
DB: 14 Gaps: 0

US-09-851-410a-6_COPY_5147_7129 (1-1983) x US-10-165-868-15 (1-327)

QY 1000 GGTGGGACGGAGCTGCGAGCTCACCACCGGCTGCTACCGCTTTTATGAGGACCTC 1059
Db 1 GlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetIysAspLeu 20

```
QY 1060 TATTTTACTAGTACTAATGGTGTGGTGGATCGCGCGGGATAGCCCTCACCCCTGTTCT 1119
Db |||||||
QY 1120 AACCTTTGCTGACACTCTGCTTGGCGGCTCGCACAGAAATTGATTTTGGTGGCTGGTGGC 1179
Db |||||||
QY 1180 CAGCTGTTCTACTCCCGTCCCGTGTCTCAGCAATGGCGAGCGGAGCTGTAAGTTGTAT 1239
Db |||||||
QY 1240 ACATCTGTAGAGAAATGCTCAGCAGGATAAGGGTATTTCGAATCCCGCATGACATTGACCTC 1299
Db |||||||
QY 1300 GGAGAATCTCGTGTGTTATTTCAGGATTATGATAACCAACATGACCAAGATCGCCGACG 1359
Db |||||||
QY 1360 CCTTCTCCAGCCCATCGCGCCCTTCTCTGCTTCAGCTAATGATGCTTGGCTC 1419
Db |||||||
QY 1420 TCTCTCAGCCCTCGCGAGTATGACCACTTATGCTCTTCCAGTGGCTCCGAGTGGTGGTGTAT 1479
Db |||||||
QY 1480 GTTCTGACTCTGTGACCTTGGTTAAATGTTGCGACCGCGCGGCGGCTGGTGGCTGG 1539
Db |||||||
QY 1540 CTGATTGGACAGGTGCACACTTGAGTGGCGCCCTCTCCACCATCCAGCAGTACTCG 1599
Db |||||||
QY 1600 AAGACCTTCTGCTCGCGCTCCGCGGTAAGTCTCTTCTGGGAGGCGAGGCACAACT 1659
Db |||||||
QY 1660 AAGCGGGGTACCTTATATTAATTAACACACTGCTAGCAGCAACTGCTTGTGAGAAAT 1719
Db |||||||
QY 1720 GCCCGCGGCGACCGGTGCGTATTTCACACTTACACACTGCTGGGTGGTGGTGGTGGT 1779
Db |||||||
QY 1780 TCCATTTCTCGGTTGCGGTTTGTAGCCCCCACTCTCGGCTAGCATTGCTTGAGGATACC 1839
Db |||||||
QY 1840 TTGACTACCTGCGCGGCCATACCTTTTGTGATGATTTTCTGCGAGAGTGGCGGCCCTT 1899
Db |||||||
QY 1900 GGCCTTCAGGCTCGGCTTTCAGCTACTGCTGCTGAGCTTACGCGCTTACAGATGAAG 1959
Db |||||||
QY 1960 GTGGGTAAACCTCGGGAGTTG 1980
Db |||||||
QY 202 ValGlyLysThrArgGluLeu 211
```

RESULT 12

```
US-10-381-770-2
; Sequence 2, Application US/10381770
; Publication No. US20040052813A1
; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPOS
; TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
; FILE REFERENCE: IBC010037PCT
```

```
; CURRENT APPLICATION NUMBER: US/10/381,770
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 214
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-2

Alignment Scores:
Pred. No.: 5,37e-185 Length: 214
Score: 210.00 Matches: 210
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.82% Indels: 0
DB: 12 Gaps: 0
```

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-381-770-2 (1-214)

```
QY 1180 CAGCTGTTCTACTCCCGTCCCGTGTCTCAGCAATGGCGAGCGGAGCTGTTAAGTTGTAT 1239
Db |||||||
QY 1240 ACATCTGTAGAGAAATGCTCAGCAGGATAAGGGTATTTCGAATCCCGCATGACATTGACCTC 1299
Db |||||||
QY 1300 GGAGAATCTCGTGTGTTATTTCAGGATTATGATAACCAACATGACCAAGATCGCCGACG 1359
Db |||||||
QY 1360 CCTTCTCCAGCCCATCGCGCCCTTCTCTGCTTCAGCTAATGATGCTTGGCTC 1419
Db |||||||
QY 1420 TCTCTCAGCCCTCGCGAGTATGACCACTTATGCTCTTCCAGTGGCTCTTCGACTGCGCCAGTTTAT 1479
Db |||||||
QY 1480 GTTCTGACTCTGTGACCTTGGTTAAATGTTGCGACCGCGCGGCGGCTGGTGGCTGG 1539
Db |||||||
QY 1540 CTGATTGGACAGGTGCACACTTGAGTGGCGCCCTCTCCACCATCCAGCAGTACTCG 1599
Db |||||||
QY 1600 AAGACCTTCTGCTCGCGCTCCGCGGTAAGTCTCTTCTGGGAGGCGAGGCACAACT 1659
Db |||||||
QY 1660 AAGCGGGGTACCTTATATTAATTAACACACTGCTAGCAGCAACTGCTTGTGAGAAAT 1719
Db |||||||
QY 1720 GCCCGCGGCGACCGGTGCGTATTTCACACTTACACACTGCTGGGTGGTGGTGGTGGT 1779
Db |||||||
QY 1780 TCCATTTCTCGGTTGCGGTTTGTAGCCCCCACTCTCGGCTAGCATTGCTTGAGGATACC 1839
Db |||||||
QY 1840 TTGACTACCTGCGCGGCCATACCTTTTGTGATGATTTTCTGCGAGAGTGGCGGCCCTT 1899
Db |||||||
QY 1900 GGCCTTCAGGCTCGGCTTTCAGCTACTGCTGCTGAGCTTACGCGCTTACAGATGAAG 1959
Db |||||||
QY 1960 GTGGGTAAACCTCGGGAGTTG 1980
Db |||||||
QY 202 SerIleSerAlaValAlaValLeuAlaPro 211
```

RESULT 13

```
US-10-381-770-12
; Sequence 12, Application US/10381770
; Publication No. US20040052813A1
; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPO
; TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
; FILE REFERENCE: IBC010037PCT
```

; CURRENT APPLICATION NUMBER: US/10/381,770
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 232
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-12

Alignment Scores:
Pred. No.: 5,32e-185 Length: 232
Score: 210.00 Matches: 210
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.82% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-381-770-12 (1-232)

QY 1180 CAGCTGTTCTACTCCCGTCCCGTGTCTCAGCAATGGCGAGCGGACTGTAAAGTTGTAT 1239
Db 23 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 42
QY 1240 ACATCTGTAGAGATGCTCAGCAGGATAAGGTATTGCAATCCGCATGACATTGACCTC 1299
Db 43 ThrSerValGluAsnAlaGlnAspLysGlyLeuAlaIleProHisAspLeu 62
QY 1300 GGAGATCTCGTGTGTTATTCAGGATTATGATAACCAATGACCAATGACCAATGGCGGAGC 1359
Db 63 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 82
QY 1360 CCTTCTCCAGCCCCATCGCGCCCTTCTCTGCTCCAGAGTAAATGATGCTTTGGCTC 1419
Db 83 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 102
QY 1420 TCTCTCAGCGTCCGAGTATGACAGTCCACTTATGGCTCTTCGACTGGCCGAGTTAT 1479
Db 103 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerThrGlyProValTyr 122
QY 1480 GTTCTGACTCTGACCTTGTTTAAATGTTGGAGCGGCGGCGGCGGCTTCCCGGTG 1539
Db 123 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 142
QY 1540 CTCGATTGACCAAGTCACTTACAGGTCCCGCTCCCGTCCCGTCCAGCAGTACTCG 1599
Db 143 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 162
QY 1600 AAGACCTTCTTGTCTGCGCTCCCGGTAAGTCTCTTTCTGGAGGACGACCAACT 1659
Db 163 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 182
QY 1660 AAAGCGGGTACCTTATAATTATACACCACTGCTAGGACCACTGCTTCCAGAA 1719
Db 183 LysAlaGlyTyrProTyrAsnTyrAsnThrAlaSerAspGlnLeuValGluAsn 202
QY 1720 GCCCGCGGACCGGTCGCTATTTCCACTTACACCACTAGCTGCTGGTCCCGTC 1779
Db 203 AlaAlaGlyHisArgValAlaIleSerThrThrThrSerLeuGlyAlaGlyProVal 222
QY 1780 TCCATTCTCGGGTTCGGTTTTAGCCCC 1809
Db 223 SerIleSerAlaValAlaValLeuAlaPro 232

RESULT 14

US-10-381-770-4
; Sequence 4, Application US/10381770
; Publication No. US20040052813A1
; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPOS
; TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
; FILE REFERENCE: IEC010037PCT

; CURRENT APPLICATION NUMBER: US/10/381,770
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 389
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-4

Alignment Scores:
Pred. No.: 9.01e-176 Length: 389
Score: 200.00 Matches: 200
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.30% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-381-770-4 (1-389)

QY 1240 ACATCTGTAGAGATGCTCAGCAGGATAAGGTATTGCAATCCGCATGACATTGACCTC 1299
Db 143 ThrSerValGluAsnAlaGlnAspLysGlyLeuAlaIleProHisAspLeu 162
QY 1300 GGAGATCTCGTGTGTTATTCAGGATTATGATAACCAATGACCAATGACCAATGGCGGAGC 1359
Db 163 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 182
QY 1360 CCTTCTCCAGCCCCATCGCGCCCTTCTCTGCTCCAGAGTAAATGATGCTTTGGCTC 1419
Db 183 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 202
QY 1420 TCTCTCAGCGTCCGAGTATGACAGTCCACTTATGGCTCTTCGACTGGCCGAGTTAT 1479
Db 203 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerThrGlyProValTyr 222
QY 1480 GTTCTGACTCTGACCTTGTTTAAATGTTGGAGCGGCGGCGGCTTCCCGGTG 1539
Db 223 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 242
QY 1540 CTCGATTGACCAAGTCACTTACAGGTCCCGTCCCGTCCCGTCCAGCAGTACTCG 1599
Db 243 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 262
QY 1600 AAGACCTTCTTGTCTGCGCTCCCGGTAAGTCTCTTTCTGGAGGACGACCAACT 1659
Db 263 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 282
QY 1660 AAAGCGGGTACCTTATAATTATACACCACTGCTAGGACCAACTGCTTCCAGAA 1719
Db 283 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 302
QY 1720 GCCCGCGGACCGGTCGCTATTTCCACTTACACCACTAGCTGCTGGTCCCGTC 1779
Db 303 AlaAlaGlyHisArgValAlaIleSerThrThrThrSerLeuGlyAlaGlyProVal 322
QY 1780 TCCATTCTCGGGTTCGGTTTTAGCCCCCTAGCTGCTGGTCCCGTCCCGTCC 1839
Db 323 SerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuGluAspThr 342

RESULT 15
US-10-381-770-3
; Sequence 3, Application US/10381770
; Publication No. US20040052813A1
; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPOS
; TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
; FILE REFERENCE: IEC010037PCT
; CURRENT APPLICATION NUMBER: US/10/381,770
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 3
; LENGTH: 194
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-3

Alignment Scores:
  Pred. No.:      1,77e-166      Length:      194
  Score:          190.00          Matches:      190
  Percent Similarity: 100.00%      Conservative: 0
  Best Local Similarity: 100.00%      Mismatches: 0
  Query Match:      28.79%          Indels:      0
  DB:               12             Gaps:        0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-381-770-3 (1-194)

QY 1240 ACATCTGTAGAGAAATCTCAGCAGGATTAAGGGTATTGCAATCCCGCATGACATTGACCTC 1299
DB 2 ThrSerValGluAsnAlaGlnGlnAspLysGlyLeAlaIleProHisaspIleAspLeu 21

QY 1300 GGAGAAATCTCGTGTGTTATTTCAGGATTATGATAACACATGACACAGATCGCCGACG 1359
DB 22 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 41

QY 1360 CCTTCTCCAGCCCATCGCGCCCTTCTCTGTCTCTTCGAGCTAATGATGTGCTTTGGCTC 1419
DB 42 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 61

QY 1420 TCTCTCACCGCTCCCGAGTATGACCATTCACCTTATGGCTCTTTCGACTGCCCCAGTTTAT 1479
DB 62 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 81

QY 1480 GTTCTGACCTCTGTGACCTTGTTAATGTGTGGACCGCGCGCAGCCGTTGCCCGTCG 1539
DB 82 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 101

QY 1540 CTCGATTGGACCAAGTTCACACTTGACGGTCGCGCCCTCTCCACCATCCAGCAGTACTCG 1599
DB 102 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 121

QY 1600 AAGACCTTCTTCTCTCGCGCTCCGCGGTAAAGCTCTCTTCTGGAGGCGAGGCACAACT 1659
DB 122 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 141

QY 1660 AAAGCGGGGTACCTTTATATATTATACACACTGCTAGCGACCACTGCTTGTTCGAGAT 1719
DB 142 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsn 161

QY 1720 GCCCGCGGGCACCGGTCGCTATTTCACATTACACCACTAGCCTGGGTGCTGGTCCCGTC 1779
DB 162 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 181

QY 1780 TCCATTTCGCGGTGCGGTTTTCAGCCCCC 1809
DB 182 SerIleSerAlaValAlaValLeuAlaPro 191
```

Search completed: August 16, 2004, 14:29:20
Job time : 122.324 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:00:26 ; Search time 97,3236 Seconds
(without alignments)
12792.733 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129

Perfect score: 3649

Sequence: 1 ATCGCGCCCTGGCGTATTTT.....GTAAACTCGGAGTGTCTAG 1983

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 2585610

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlip
-Q=/cgm2_1/USPTO.spool_p/US09851410/runat_16082004_125455_8610/app_query.fasta_1.4238
-DB=Published Applications AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blotsum2
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09851410@cgm2_1 27 @runat_16082004_125455_8610
-NCFU=6 -ICFU=3 -NO_MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA:*
1: /cgm2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgm2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgm2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgm2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgm2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgm2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgm2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgm2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgm2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgm2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgm2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgm2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgm2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgm2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgm2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgm2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgm2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgm2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

ALIGNMENTS

RESULT 1

US-09-769-066-13
; Sequence 13, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; McAtee, C. Patrick
; Yarbrough, Patrice O.
; Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESS: Dehlinger & Associates

STREET: 350 Cambridge Ave., Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

Sequence 13, Appl
Sequence 8, Appl
Sequence 19, Appl
Sequence 1, Appl
Sequence 14, Appl
Sequence 20, Appl
Sequence 57, Appl
Sequence 51, Appl
Sequence 55, Appl
Sequence 92, Appl
Sequence 92, Appl
Sequence 167, App
Sequence 167, App
Sequence 15, Appl
Sequence 25, Appl
Sequence 27, Appl
Sequence 16, Appl
Sequence 26, Appl
Sequence 28, Appl
Sequence 17, Appl
Sequence 18, Appl
Sequence 206, App
Sequence 207, App
Sequence 206, App
Sequence 207, App
Sequence 4, Appl
Sequence 17, Appl
Sequence 15, Appl
Sequence 199, App
Sequence 200, App
Sequence 199, App
Sequence 200, App
Sequence 18, Appl
Sequence 16, App
Sequence 176, App
Sequence 175, App
Sequence 175, App
Sequence 189, App
Sequence 189, App
Sequence 6, Appl
Sequence 12, Appl
Sequence 2, Appl


```
/ APPLICATION NUMBER: US/09/769,066
/ FILING DATE: 24-Jan-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/542,634
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0293.30
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: Hepatitis E Virus (Burma strain)
/ ORF-2
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-769-066-13

Alignment Scores:
Pred. No.: 5,91e-243 Length: 660
Score: 3408.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.40% Indels: 0
DB: 9 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-769-066-13 (1-660)

QY 1 ATGCGCCCTCGGCCTATTTGTTGCTGCTCTCATGTTTTCCTATGTCGCCCGCGCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGGCCGGTACGCGTCTGGCGCGCTGTCGGCGCGCGAGCGGGTTCGGCGGTGT 120
DB 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGly 40
QY 121 TTCTCGGGTGACGGGTTCATTCTAGCCCTTCGCAATCCCTATATTCATCCCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProIlyrIleHisProThrAsn 60
QY 181 CCTTCGCCCGCGATGTCACCGCTGCGCGCGGGCTGGACCTCGTGTTCGCAACCGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGGCTGGCGTGCACGAGCCCGCGCGCGCGCTGCTCAGTGT 300
DB 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACCTACACAGTGGGGCCCGCGCTAACCGCGTGCCTCCGCGCCCATCACACCCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAATGCTGATGTGACCTCCGCGCGCGCATCTTCGCGCGCGCAGTATACCATCAACA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTrpAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCGCGCGCGCATACCTGGTTCCTTATGCGCGCGCTCT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGCGCTTTACCTCTTCGCGCGCGCGCATACCTGGTTCCTTATGCGCGCGCTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTGCCGTGCCACAAATCCGTTACCGCGCGCTGTCGCCAAT 600

181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
601 GCTGTGCGGGTTAGCCATCTCCATCTCATCTCTGCGCACACACACACACCGCGAG 660
201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
661 TCCGTTGATATGAATTCATACCTACCTGAGGATGTTTCGTTATTTAGTCCAGCCGCGCAT 720
221 SerValAspMetAsnSerIleThrSerThrAspValArgIleuValGlnProGlyIle 240
721 GCTCTGAGCTTGTGATCCCAAGTACGCGCTACCTATCGTAACCAAGGCTGGCGCTCC 780
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
841 CATGCTCCTCACTGTAATTCCTATATACACCTATACACCTGCGGTCCTCGGGCTGTTG 900
281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
901 GACTTTGCCCTTACGCTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC 960
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
961 TCCGTTATTCAGCACTGCTCGCCACCGCTTCGCGGGTGGCGGAGGAGGAGGAGGAG 1020
321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
1021 CTCACCAACCGCTGCTACCGCTTATGAAGGAGCCTCTATTTACTAGTACTAATGCT 1080
341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
1081 GTCGCTGAGATCGCGCGGGATAGCCCTACCCCTGTTCAAACCTGCTGACACTCTGCTT 1140
361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
1141 GCGCGCTCGCGACAGAAATGTTGCTGCGCTGCTGCGCGCGCGCTGCTTCTACTCCGCTCC 1200
381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
1201 GTTGTCTCAGCAATCGCGGAGCGGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
1261 CAGGATAAGGTTATTCGAATCCCGCATGACATGACCTCGGAGAAATCTCGTGTGTTATT 1320
421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle 440
1321 CAGGATTATGATAACCAACATCAAGATCGCGCGCGCGCTTCTCCAGCCCGCATCGCG 1380
441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
1381 CCTTCTCTGCTCCTTCGAGCTAATGATGCTTTCGCTTTCGCTTCTCTCACCGCTCCGAGTAT 1440
461 ProPheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyr 480
1441 GACCACTGACATTATGGCTCTTCGAGTGGCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500
481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
1501 GTTATGTTGGACCGCGCGCGCGCGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
501 ValAsnValAlaThrGlyAlaGlnAlaValAlaAlaSerLeuAspTrpThrLysValThr 520
1561 CTTGAGCGTCCGCCCTCTCCACCATCCAGCAGTACTGAAGACCTCTTTTCTCTGCGCG 1620
521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
1621 CTCGCGGTAAAGTCTCTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
```

```
Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGCTAGCGACCACTGCTCTCCAGAGATCGCGGGACCGGGTGGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCACACTTACACACTAGCTGGGTGCTGCTCCGCTCTCCATTTCTTCGGGTGGCGTT 1800
Db 581 IleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTGCGCTAGCACTGCTTGGAGTACTTGGACTACCTGCGCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGATGATTTCTGCCAGAGTGGCGCGCCCTTGGCTTCCAGGGCTGGCTTTC 1920
Db 621 HisThrPheAspPheCysProGlnCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTGAGCTTTCAGCGCCTTAAGATGAAGTGGGTAAAACTCGGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 2
US-09-851-410-8
; Sequence 8, Application US/09851410
; Publication No. US20030124510A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; Yarbough, Patrice O
; Bradley, Daniel W
; Krawczynski, Krzysztof Z
; Tam, Albert
; FTY, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 07-May-2001
; PRIORITY DATE: 07-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/128,275
; FILING DATE: <unknown>
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
```

```
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-851-410-8
Alignment Scores:
Pred. No.: 5,91e-243 Length: 660
Score: 3408.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.40% Indels: 0
DB: 10 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-851-410-8 (1-660)
QY 1 ATGCGCCCTCGGCCTATTGTTGCTGCTCATGTTTTCCTATGCTGCTGCGCGCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGGTACAGCGCTCTGGCGCGCGTGTGGCGGGCGCGACGGCGGTTCGGGGTGGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlySerGlyGlyGly 40
QY 121 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCTTCGCGCCCGATGTCACCGCTGCGCGCGGGCTGACCTCGTGTTCGCAACCCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCTCGCTTGGCGTGACCGCGCCAGCGCCCGCGGTTCGCTCAGCTGT 300
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACTACACAGCTGGGGCGCGCGCTAACCGGGTGTGCTCGCGCGCATGACACCCCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGTGATGTGACTCCCGCGCGCCCATCTTGGCGGGAGTATTAACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCCGTGGCCACCGGCACCTAACCTGGTTCCTTTATCGCGCCCTCT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGGCTTTTACCCCTTCAGGACGGCACCAATACCCATATAATGGCCACGAGACTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTGCCCGTGCACAAATCCGTTACCGCGCTGGTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCCGGGTACGGCATCTCCATCTCATTTCTGGCCACAGACACACACCCCGCAGC 660
Db 201 AlaValGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATTAACCTCAGCGAGTTCGTATTTTAGTCCAGCGCGGATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTGTGATCCCAAGTGGCGCTACACTATCGTAACCAAGCTGGCGTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260
QY 781 GTCCAGACCTCTCGGGTGGCTGAGGAGGAGCTACCTCTGCTCTGTTATGCTTTGGATA 840
```

Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysile 280
Qy 841 CATGGCTCACTCGTAAATCCCTATCTACTAATACACCTATACCGGTGCGCTCGGGCTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
Qy 901 GACTTTGGCTTGGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACGGGGTTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
Qy 961 TCCGTTATTCAGACACTGCTCGCACCGCTTCGCGGTGCGGAGCGGAGCTCGGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
Qy 1021 CTCACACACCGCTGCTACCGCTTATGAGGACCTCTATTCTAGTACTAATGGT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
Qy 1081 CTCGCTGAGATCGCGCGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTT 1140
Db 361 ValGlyGluLeuGlyArgGlyLeuAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
Qy 1141 GCGCGCTGCGCACAGATTGATTCGTCGCTGCTGCGTGGCCAGCTGTTCTACCTCCGTC 1200
Db 381 GlyGlyLeuProThrGluLeuLeuSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
Qy 1201 GTTGTCTCAGCAATGGCGGAGCGGACTGTTAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
Qy 1261 CAGGATAAGGGTATTGCAATCCGATCCGATGACATTCGCGAGAAATCTGCTGTGGTTAT 1320
Db 421 GlnAspLysGlyLeuAlaIleProHisAspLeuLeuGlyGluSerArgValValle 440
Qy 1321 CAGGATTATGATAACACATGAACAGATCGGCGAGCCCTTCTCCAGCCCCATCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
Qy 1381 CCTTTCTCTGCTCGAGCTAATGATGTCTTGGCTCTCTCTCACCGCTCCCGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuThrLeuSerLeuThrAlaAlaGluTyr 480
Qy 1441 GACCACTCCACTATGGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
Qy 1501 GTTAATGTTGCAACGGCGCGAGCGCTTGGCGGTGCGTCTGATGGACGAGGTACA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
Qy 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTGCTGCGG 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
Qy 1621 CTCGCGGTAACTCTCTTTCTGGAGGAGCAGCAACACTAAAGCCGGGTACCTTATAAT 1680
Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrLysAlaGlyTyrProTyrAsn 560
Qy 1681 TATACACACTGCTAGGACCAACTGTTGTCGAGATGCGCGGCGGCGGCGGCTGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
Qy 1741 ATTTCCACTTACACACTAGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
Qy 1801 TTAGCCCCCACTCTGCGGTAGCATGCTTGGAGGATACCTTGAGTACCTCGACTACCTCGCGCGC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
Qy 1861 CATACTTTGATGATTTGCGCCAGAGTGGCGCCCTTGGCTTGGCTTGGCTTGGCTTGGCTT 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640

Qy 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAGATGAAGGTGGGTAAACTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660
RESULT 3
US-10-165-868-19
; Sequence 19, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; Bradley, Daniel W.
; Twu, Jr-Shin
; Purdy, Michael A.
; Tam, Albert W.
; Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/165,868
; FILING DATE: 06-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-165-868-19
Alignment Scores:
Pred. No.: 5,91e-243 Length: 660
Score: 3408.00 Matches: 660
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0

; CURRENT APPLICATION NUMBER: US/10/381,770
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 660
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-1

Alignment Scores:

Pred. No.: 1,066-241 Length: 660
Score: 3391.00 Matches: 656
Percent Similarity: 99.55% Conservative: 1
Best Local Similarity: 99.39% Mismatches: 3
Query Match: 92.93% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-381-770-1 (1-660)

QY 1 ATGCGCCTCGCCCTATTGCTGCTGCTCTCTCATGTTTTCCTATGCTGCCGCGCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuLeuMetPheLeuProMetLeuProIlePro 20
QY 61 CCGCCGGTCAAGCCTCTGCGCGCGCTGCTGCGCGCGCAGCGCGGTTCCGCGCGTGGT 120
DB 21 ProProGlyGlnProSerGlyArgArgArgGlyArgGserGlyGlySerGlyGlyGly 40
QY 121 TTCTGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTrpIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCGATGTCACCGCTCGCGCGCGGCTGGACCTCGGTGTCGCCAACCCGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCGCTGGCGTGACAGGCCAGCGCCCGCGCTTGCCTCAGCTCGT 300
DB 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaAlaAlaSerArgArg 100
QY 301 AGACCTACAGCTGGGCGCGCGCTAACCGCGGTGCTCCGCGCGCATGACACCGCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTCAGCTCCCGCGCGCACTTTCGCGCGCAGTATAACCTATCAACA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTrpAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCCGTGGCCACCGCACTAACCTGGTTCTTTATGCGCGCCCTCTT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGCTTTTACCTTCAGACGCGCACCAATACCAATATAATGCGCACGGAAGCTTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGTTCCCGTGCCCAATCCGTTACCGCGCGCTGGTCCCCCAAT 600
DB 191 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGCGGTACGCCATCTCCATCTCATTCGTGCCACAGACCAACCAACCCCGACG 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
QY 661 TCCGTTGATATCAATCAATACCTCGAGGATGTCGTTATTTAGTCCAGCCCGGCATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGATCCCAAGTAGCGCTTACACTATCGTAACCAAGGCTGGCGCTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisIstYrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTGTTATGCTTTGCATA 840

DB 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCAGTCGTAAATTCCTATACCTATACACCTATACCGGTGCTCGGCTGTG 900
DB 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCGCTTGTAGCTTGTGCAACCTTACCCCGGTAAACACCAATACCGGGTGC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGTTTATCCAGCACTGCTCGCACCGCTTCTGTCGCGTGCAGACGGAGCTCCGAG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCGCTGCTACCGCTTATGAGGACCTCTATTTTACTAGTACTAATGCT 1080
DB 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTTGTGCTGACACTCTGCTT 1140
DB 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTGCGCAGACAAATTGATTTCGCGCTGGTGGCCAGCTGTTCTATCTCCGCTCC 1200
DB 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTCTCTCAGCAATGCGGAGCGGCTGTTAAGTTGTATACATCTGTAGAGAACTGCTCAG 1260
DB 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGATAAGGTATTGCAATCCCGCATGACATTCACCTCGGAGAACTCTGTTGCTTATT 1320
DB 421 GlnAspLysGlyIleAlaIleProHisAspLeuLeuGlyGluSerArgValValIle 440
QY 1321 CAGATTATGATAACAAACATGAACAGATCGCGCAGCGCTTCTCCAGCCCATCGGCG 1380
DB 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTCTCTGCTTCGAGCTAATGATGCTGTTTGGCTCTCTCACCGCTGCCGAGTAT 1440
DB 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTCTGACTCTGTGACCTTG 1500
DB 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGGACCGCGCGCGAGCCGTTGCGCGGTGCTCGATTGACCAAGGTACA 1560
DB 501 ValAsnValAlaThrGlyAlaGlnAlaValaAlaArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTTCGAGCTGCCCCCTCTCCACCATCCAGCAGTACTCGAAGACCTCTTTGTCTCTGCGG 1620
DB 521 LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAAAGCTCTCTTCTGGAGGACGACACAACTAAAGCCGGGTACCTTTATAAT 1680
DB 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGTAGCGACCAACTGCTTTCGGAATGCGCGCGGACCGGCTCGCT 1740
DB 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACCACCTAGCTGGTGTGCTGCCGCTCTCATTTCTGCGGTTCGCGT 1800
DB 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCTGCTTGGAGTACTGCTGAGTACCTCGCCCGCGCC 1860
DB 601 LeuAlaProHisSerValLeuAlaLeuLeuGluAspThrMetAspTyrProAlaArgAla 620
QY 1861 CATACTTTGATGATTTCTGCGCAGTGCGCCCTTGGCTTACGCTTCAGGCTCGGCTTTC 1920

Qy 841 CATGGCTCACTCGTAATTCCTATACTAAATACACCTATACCGGTGCCCTCGGCTGTG 900
Db |||||||
Qy 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
Db |||||||
Qy 901 GACTTGGCTTGAAGTTGAGTTTCGGAACCTTACCCCGGTAAACACCAATACCGGGTCT 960
Db |||||||
Qy 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
Db |||||||
Qy 961 TCCCGTTATCCAGCACTGCTCGCACCGCTTCTCGTGGCGGTGGAGCGGACGCGCGAG 1020
Db |||||||
Qy 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
Db |||||||
Qy 1021 CTCACCAACCGCTGTACCCGCTTATGAAGACCTCTATTTTACTAGTACTAATGGT 1080
Db |||||||
Qy 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGly 360
Db |||||||
Qy 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACTCTGTTCAACCTTGTGACACTCTGCTT 1140
Db |||||||
Qy 361 ValGlyGluValGlyArgGlyLeuAlaLeuThrLeuLeuAenLeuAlaAspThrLeuLeu 380
Db |||||||
Qy 1141 GCGGCTGCGCGACAGAAATGATTTCGTGGCTGGTGGCCAGCTTCTACTCCGCTCC 1200
Db |||||||
Qy 381 GlyGlyLeuProThrGluLeuLeuSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
Db |||||||
Qy 1201 GTTGCTCTCAGCAATGGCGGACCGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
Db |||||||
Qy 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
Db |||||||
Qy 1261 CAGATTAAGGTATTGCAATCCCGCATGACANTGACCTCGGAGAAATCTCGTGGTTATT 1320
Db |||||||
Qy 421 GlnAspLysGlyValAlaLeuProHisAspLeuAspLeuGlyAspSerArgValValLe 440
Db |||||||
Qy 1321 CAGATTAATGATACCAACATGACACAGATCGCGGCGCTCTCCAGCCCACTCGGC 1380
Db |||||||
Qy 441 GlnAspTyrAspAsnGlnHisGlnAlaAspArgProThrProSerProAlaProSerArg 460
Db |||||||
Qy 1381 CCTTCTCTGCTTCGAGCTTAATGATGCTGTTGGCTCTCTCTCACCGCTCGCGAGTAT 1440
Db |||||||
Qy 461 ProPheSerValLeuArgAlaAsnAspValLeuTyrPleuSerLeuThrAlaAlaGluTyr 480
Db |||||||
Qy 1441 GACAGTCCACTTATGCTCTTCACTGGCGGCACTTATGTTTCTGACTCTGTGACCTTG 1500
Db |||||||
Qy 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLeu 500
Db |||||||
Qy 1501 GTTAATGTTGCGCGCGCGCGCTGCGCGCTGCTCGCTCGATGCGACCAAGGTGACA 1560
Db |||||||
Qy 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrPheLysValThr 520
Db |||||||
Qy 1561 CTTGACGCTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTGCGCG 1620
Db |||||||
Qy 521 LeuAspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuPro 540
Db |||||||
Qy 1621 CTCGCGGTGAAGCTCTCTTCTGGGAGCAGGCACTAAAGCGGGTACCCCTTATAAT 1680
Db |||||||
Qy 541 LeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
Db |||||||
Qy 1681 TATAACACCACTGTAGGACCACTGCTTCGAGATCGCGGCGGCGGCGGCTGCT 1740
Db |||||||
Qy 561 TyrAsnThrThrAlaSerAspGlnIleLeuLeuGluAsnAlaAlaGlyHisArgValAla 580
Db |||||||
Qy 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCGCTCTCCATTTCTCGGTTGCGGTT 1800
Db |||||||
Qy 581 IleSerThrTyrThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaVal 600
Db |||||||
Qy 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGGACTACCCCTGCGCGGCC 1860
Db |||||||
Qy 601 LeuAlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAla 620
Db |||||||
Qy 1861 CATACTTTTGTATTTCTGCCAGAGTGGCGCCCTTGGCTTACGGGTGCGCTTTC 1920
Db |||||||
Qy 621 HisThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPhe 640
Db |||||||
Qy 1921 CAGTCTACTGTGCTGAGCTTTCAGCGCCTTAAGATGAAGTGGGTAAACTCGGAGTTG 1980
Db |||||||

Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysValGlyLysThrArgGluLeu 660
|||
RESULT 7
US-10-165-868-20
; Sequence 20, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; Bradley, Daniel W.
; Twu, Jr-Shin
; Purdy, Michael A.
; Tam, Albert W.
; Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/165,868
; FILING DATE: 06-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 9
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-10-165-868-20
Alignment Scores:
Pred. No.: 3,45e-227 Length: 659
Score: 3194.50 Matches: 614
Percent Similarity: 96.21% Conservative: 21
Best Local Similarity: 93.03% Mismatches: 24
Query Match: 87.54% Indels: 1
DB: 14 Gaps: 1

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-165-868-20 (1-659)

QY 1 ATGCGCCCTCGCCCTATTGTTGCTGCTCCTCATGTTTTCCTATGTCGCCGGCCA 60
Db 1 MetArgProArgProLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY 61 CGCGCCGGTCAGCGCTCTGGCCGCGTCTGGTGGCGCGCAGCGGGGTTCCGGCGGTGT 120
Db 21 ProThrGlyInProSerGlyArgArgArgGlyArgGserGlyGlyThrGlyGlyGly 40
QY 121 TTCTGGGGTACCGGGTGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCATGTACCGCTCGCGCGGGGTGGACCTGCTGTTCGCCAACCGGC 240
Db 61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTTGGCGTGACACAGGCCCGCGCGCGCTTGCCTCACGTGCT 300
Db 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
QY 301 AGACCTACACAGCTGGGCGCGCGCTAACCGCGTGTCTCCGCGCCCATCACCCCG 360
Db 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCAGTCTGATGTGATCCCGCGCGCCATCTTGGCGCGCAGTATACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGTCGCCACCGCACTAACCTGGTCTTTATGCCCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGCGTTTACCCTTCAGACCGCACCAATACCATATATATGCGCACGGAAGCTTCT 540
Db 161 AsnProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGTTCCCGTCCCAATCCGTTACCGCGCTGTCTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GTGTGCGCGGTAGCCATCTCCATCTCATCTGCGCACACACACACACCCCGCAG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATGAATCAATTAACCTCGACGATGTTGCTATTTTATGTCAGCCCGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGATCCCAAGTCAGCGCCTACACTATCGTAACCAAGGCTGGCGTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCCAGACCTCTGGGGTGGCTGAGGAGGCTACTCTGTGCTTGTATGCTTTGCGATA 840
Db 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCAGTCGTAAATTCCTATTAATAACCCCTATACCGGTACCGGTGCCCTGGGTGTG 900
Db 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGGCTTGAGCTTGAGTTTCGACCTTACCCCGGTAAACACCAATACGGGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCAGCACTGCTCGCACCGCCCTTCGTCGGGTGCGGACGGGACTCCCGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisSer---AlaArgGlyAlaAspGlyThrAlaGlu 339
QY 1021 CTCACACACGGCTGCTACCGCTTTATGAAGGACCTCTATTTACTAGTACTAATGGT 1080
|||||

Db 340 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGly 359
QY 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCTGTTCAACCTTGTCTGACACTCTGCTT 1140
Db 360 ValGlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeu 379
QY 1141 GCGCGCCTCGCGACAGAATTGATTTCGTCGGTGTGGCCAGCTGTTCTACTCCCGTCCC 1200
Db 380 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 399
QY 1201 GTTGTCTCAGCAATGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db 400 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 419
QY 1261 CAGGATAGGGTATTGCAATCCGCATCCGATGACATTGACCTCGGAGAACTCGTGTGGTATT 1320
Db 420 GlnAspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValIle 439
QY 1321 CAGGATTATGATAACCAACATGAAACAGATCGCGCAGCGCTTCTCCAGCCCATCGCGC 1380
Db 440 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 459
QY 1381 CTTTCTCTGCTTCGAGCTTAATGATGTGCTTTCGCTCTCTCTCACCGCTCCGAGTAT 1440
Db 460 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 479
QY 1441 GACCACTCCACTTATGGCTCTTCGACTGGCCGAGTTATGTTTCTGACTCTGACCTG 1500
Db 480 AspGlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLeu 499
QY 1501 GTTAATGTTGGACCGCGCGCAGCGCTTGGCCGCTGCTCGATGAGCAAGGTCAACA 1560
Db 500 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr 519
QY 1561 CTTGACGGTGGCCCTCTCCACCATCCAGCAGTACTGAGAACCTTCTTTGCTCTCGCG 1620
Db 520 LeuAspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuPro 539
QY 1621 CTCGCGGTAGCTCTCTTCTGGGAGGAGCAGCACAACTAAAGCCGGTACCCTTATAT 1680
Db 540 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 559
QY 1681 TATAACACCACTGTAGCGACCACTGCTGTGCGAATGCGCGCGGCGCACCGGTCGT 1740
Db 560 TyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAla 579
QY 1741 ATTTCACCTTACACACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Db 580 IleSerThrTyrThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaVal 599
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTGCTTGAGGATACCTTGACTACCTCCCGCGCGC 1860
Db 600 LeuAlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAla 619
QY 1861 CATACTTTGATGATTCTGCCAGAGTCCCGCCCCCTTGGCCTTTCAGGGCTCGCTTTC 1920
Db 620 HisThrPheAspAspPheCysProGluCysArgAlaLeuLeuGlyLeuGlnGlyCysAlaPhe 639
QY 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCCTTAAGATGAAGTGGGTAAACTCGGGAGTTG 1980
Db 640 GlnSerThrValAlaGluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 659

RESULT 8

US-10-239-090A-57

; Sequence 57, Application US/10239090A

; Publication No. US20040101820A1

; GENERAL INFORMATION:

; APPLICANT: KABUSHIKI KAISHA TOSHIBA

; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co
; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same,
; TITLE OF INVENTION: detecting hepatitis E virus using the same

; FILE REFERENCE: 02S0741P

; CURRENT APPLICATION NUMBER: US/10/239,090A

```
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 57
; SEQ ID NO 57
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Hepatitis E Virus JAK-Sai (ORF2)
US-10-239-090A-57

Alignment Scores:
Pred. No.: 4.84e-226      Length: 674
Score: 3179.00           Matches: 607
Percent Similarity: 95.90% Conservative: 25
Best local Similarity: 92.11% Mismatches: 27
Query Match: 87.12%      Indels: 0
DB: 16                   Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-239-090A-57 (1-674)

QY 1 ATGCGCCTCGGCTATTGTTGCTGCTCCTCATGTTTTTGGCTATGCTGCGCGGCCA 60
DB 15 MetArgSerArgAlaLeuPheLeuPheLeuPheValPheLeuProMetLeuProAlaPro 34
QY 61 CCGCCCGGTACGCGCTCTGGCCGCGTCTGGGGGCGGCGACGCGCGGTTCGCGCGTGGT 120
DB 35 ProAlaGlyGlnProSerGlyArgArgArgGlyArgSerGlyGlyAlaGlyGlyGly 54
QY 121 TTCTGGGGTGACCGGGTGTATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC 180
DB 55 PheTrpGlyAspArgValAspSerGlnProPheAlaLeuProTyrIleHisProThrAsn 74
QY 181 CCCTTCGCCCGCGATGTCACCGCTCGCGCGCGGCTGACCTGCTGTTCGCGACCGGCC 240
DB 75 ProPheAlaSerAspIleProThrAlaAlaGlyAlaGlyAlaArgProArgGlnProAla 94
QY 241 CGACCACTCGGCTCGCTTGGCGGTGACACGAGCCGCGCGCGCGCTGTTCCTCAGTCGT 300
DB 95 ArgProLeuGlySerAlaTrpArgAspGlnSerGlnArgProThrAlaSerAlaArgArg 114
QY 301 AGACCTACCAAGTGGGGCGCGCGGTAAACCGCGGTCTCGCGGCCATGACACCCCG 360
DB 115 ArgProAlaProAlaGlyAlaAlaProLeuThrAlaValAlaProAlaProAspThrAla 134
QY 361 CCAGTGCCTGATGTCGACTCCCGCGGCGCATCTTGGCGGCGAGTATACCTATCAACA 420
DB 135 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 154
QY 421 TCTCCCTTACCTCTTCGCTGGCCACCGGCACCTAACCTGCTTCTTTATGCGCGCCCTCT 480
DB 155 SerProLeuThrSerThrIleAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 174
QY 481 AGTCGCGTTTACCCCTTCAGGACGACCAATACCCTATATATGCGCAGGAAGCTTCT 540
DB 175 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 194
QY 541 AATTTATGCCAGTACCGGGTTCGCGTCCCAATCCGTTACCGCGCGTGGTCCCAAT 600
DB 195 AsnTyrAlaGlnTyrArgValValArgAlaThrIleArgTyrArgProLeuValProAsn 214
QY 601 GCTGTCGCGGGTTACGCGCATCTCCATCTCATCTTGGCCACAGACCAACACCCCGCAGC 660
DB 215 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 234
QY 661 TCGCTTGAATGAATCAATAACCTCGACGGATGTTCTGATTTTATGTCAGCCGGCATA 720
DB 235 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 254
QY 721 GCCTCTCAGCTGTGATCCCAAGTGAGCGCTACACTATCGTACCAAGGCTGCGCTCC 780
DB 255 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTTrpArgSer 274
QY 781 GTCGAGACCTCTGCGGGTGGCTGAGGAGGAGGCTACCTCTGCTCTTGTATGCTTTGCATA 840
```

```
DB 275 ValGlnThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 294
QY 841 CATGCTCATCTCGTAATAATTCCTATACATACACCTATACCGGTGCCTCGGGCTGTG 900
DB 295 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 314
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTACACCAATACCGGGGTC 960
DB 315 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 334
QY 961 TCCCGTTATTCCAGCACTGCTCGCCACCGCTTCGTCGCGGTGGGACGGGACTGCCGAG 1020
DB 335 SerArgTyrSerSerSerAlaArgHisIlyLeuArgGlyProAspGlyThrAlaGlu 354
QY 1021 CTACACACCGGCTGCTACCGCTTATGAGAGACCTCTATTTTACTAGTACTAATGGT 1080
DB 355 LeuThrThrThrAlaAlaThrArgPheMetIlyAspLeuHisPheThrGlyThrAsnGly 374
QY 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACCTCGCTT 1140
DB 375 ValGlyGluValGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 394
QY 1141 GCGCGCTCGCCGACAGAAATTGATTTCGTCGCGTGGTGGCCAGCTGTTCCTACTCCCGTCCC 1200
DB 395 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 414
QY 1201 GTTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
DB 415 ValValSerAlaAsnGlyGluProThrValIlyLeuTyrThrSerValGluAsnAlaGln 434
QY 1261 CAGGATAAGGGTATTGCAATCCGCAATCGACCTCGGAGAACTCTCGTGTGGTTATT 1320
DB 435 GlnAspIysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 454
QY 1321 CAGGATTATGATAACCAACATGAACAAGATCGGCGCGCTTCCTCCAGCCCAAGTCCGCGC 1380
DB 455 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 474
QY 1381 CTTTCTCTGCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCAGCGCTGCGAGTAT 1440
DB 475 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 494
QY 1441 GACCAGTCCACTTATGCTCTTCGACTGCGCGCTTATGTTTATGCTGCTGACCTCTG 1500
DB 495 AspGlnThrThrTyrGlySerSerThrAsnProMetTyrValSerAspThrValThrPhe 514
QY 1501 GTTAATGTTCCGACCGCGCGCGCTTGCCTGCGCTCGATTGCGACCAAGTCCACA 1560
DB 515 ValAsnValAlaThrGlyAlaGlnGlyValSerArgSerLeuAspTrpSerIlyValThr 534
QY 1561 CTTGACGGTCCGCCCTCTCCACCATCCAGAGTACTCGAAGACCTTCTTTGCTCCTGCGC 1620
DB 535 LeuAspGlyArgProLeuThrThrIleGlnGlnTyrSerIlyThrPhePheValLeuPro 554
QY 1621 CTCGCGGTAAAGCTCTCTTCTCGGGGAGGACACAACTAAAGCCGGGTACCTTATAAT 1680
DB 555 LeuArgGlyIysLeuSerPheTrpGluAlaGlyThrThrIlyAlaGlyTyrProTyrAsn 574
QY 1681 TATAACCACTGCTAGCGACCAACTGCTTGTGCGAATCCGCGCGGCGCACCGGCTCGCT 1740
DB 575 TyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValCys 594
QY 1741 ATTTCCACTTACACCTAGCCTAGCTGGGTGCTGGTCCGCTCTCCATTTCTCGGTTGCCGTT 1800
DB 595 IleSerThrTyrThrThrAsnLeuGlySerGlyProValSerIleSerAlaValGlyVal 614
QY 1801 TTAGCCCCCAGCTCGGCTAGCATTCGTTGAGGATACCTTGGACTACCTGCGCGCGGCC 1860
DB 615 LeuAlaProHisSerAlaLeuAlaLeuGluAspThrValAspTyrProAlaArgAla 634
QY 1861 CATACTTTTTCATGATTTCTGCCAGAGTGCAGCGCCCTTGGCTTCAAGGCTCGGCTTTC 1920
```

Db 635 HisThrPheAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe 654
Qy 1921 CAGTCTACTGCTGCGCTGAGCTTCAGCGCTTAAGATGAAGTGGGTAAACTCGGAG 1977
Db 655 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu 673
RESULT 9
US-10-239-090A-51
; Sequence 51, Application US/10239090A
; Publication No. US20040101820A1
; GENERAL INFORMATION:
; APPLICANT: KABUSHIKI KAISHA TOSHIBA
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co
; TITLE OF INVENTION: from Japanese chip including the same, kit including the same,
; TITLE OF INVENTION: detecting hepatitis E virus using the same
; FILE REFERENCE: 0280741P
; CURRENT APPLICATION NUMBER: US/10/239,090A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 57
; SEQ ID NO 51
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus JKN-Sap (ORF2)
US-10-239-090A-51
Alignment Scores:
Pred. No.: 8,03e-226 Length: 660
Score: 3176.00 Matches: 607
Percent Similarity: 96.05% Conservative: 26
Best Local Similarity: 92.11% Mismatches: 26
Query Match: 87.04% Indels: 0
Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-239-090A-51 (1-660)
Qy 1 ATGCGCCCTCGCCCTATTGTTGCTGCTCCCTCANGTTTTTGCCTATGCTGCCCGGCCA 60
Db 1 MetArgProArgAlaValLeuLeuLeuPheLeuLeuValLeuLeuProMetLeuProAlaPro 20
Qy 61 CGCGCCGCTGACCGCTCTGCGCGCTGCTGGCGCGCGCGCGCGCTGCGCGCGTGT 120
Db 21 ProAlaGlyGlnProSerGlyArgArgGlyArgGlyArgSerGlyGlyAlaGlyGly 40
Qy 121 TTCTGGGTGACCGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaLeuProTyrIleHisProThrAsn 60
Qy 181 CCTTTCGCCCGCGATGTACCCCTCGCGCGCGCGCTGGACCTCGTGTTCGCCAACCCGCG 240
Db 61 ProPheAlaAlaAspValValSerGlnProGlyAlaGlyThrArgProArgGlnProPro 80
Qy 241 CGACCACTCGCTCGCTGCGCTGACAGAGCCCGCGCGCGCGCTGCGCTGCGCTGCT 300
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnSerGlnArgProSerAlaAlaProArg 100
Qy 301 AGACCTACCACTGCGCGCGCGCTAACCGCGCTGCTGCGCGCGCGCTGCGCGCGCG 360
Db 101 ArgSerAlaProAlaGlyAlaAlaProLeuThrAlaValSerProAlaProAspThrAla 120
Qy 361 CCAGTGCCTGATGCTGCGCTGCGCGCGCGCTGCGCGCGCGCTGCGCGCGCTGCGCG 420
Db 121 ProValProAspValAspSerArgGlyAlaAlaLeuArgGlnGlnTyrAsnLeuSerThr 140
Qy 421 TCTCCGCTTACTCTTCCGTGCGCACCGGCACTAACCTGGTCTTTATGCGCGCGCTCT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
Qy 481 AGTCCGCTTTTACCCCTTCAGGACCGCACCAATACCATATATATGCGCACGAGCTTCT 540
Db 161 AsnProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180

Qy 541 AATTATGCCAGTACCGGTTGCCCGTCCCAATCCGTTACCGCCCGCTGGTCCCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValValAlaThrIleArgTyrArgProLeuValProAsn 200
Qy 601 GCTGTCGGCGGTACGCCATCTCCATCTCATCTCTGGCCACAGCACACACCCCGGAG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
Qy 661 TCCGTTGATATGAATTCAATACCTCGACGAGTGTTCGTTATTTAGTCCAGCCCGGCATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
Qy 721 GCCTCTGAGCTTGATCCCAAGTCAGCGCTTACACTATCGTAACCAAGGCTGGCGTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
Qy 781 GTCGAGACCTCTGGGTGGCTGAGGAGGCTACTCTGGTCTTGTATGTTTGCATAT 840
Db 261 ValGluThrThrGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
Qy 841 CATGGCTCACCTCGTAAATTCCTATATAATACACCTTATACCGTGCCTCGGGCTGTG 900
Db 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
Qy 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
Qy 961 TCCGCTTATTCAGCACTGCTCGCACCGCTTCTGCGGTGCGGAGCGGAGCTGCCGAG 1020
Db 321 SerArgTyrThrSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
Qy 1021 CTCACCCACCGCTGCTACCGCTTATGAGGACCTCTATTTTACTAGTACTAATGCT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyThrAsnGly 360
Qy 1081 CTCGCTGAGATCGCGCGGGATAGCCCTCACCCCTGTTCAACCTTCTGACACTCTGCT 1140
Db 361 ValGlyGluValGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
Qy 1141 GCGCGCTTCCGACAGAAATTGTTTCGCGGTGCTGGCCAGCTGTTCTACTCCCGTCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
Qy 1201 GTTGCTCAGCCCAATGGCGAGCCGACTGTAAAGTTGTATACATCTGTAGAGAACTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
Qy 1261 CAGGATAAGGTPATTGCAATCCCGCATGACATTCACCTCGGAGAACTCTGCTGTGTTAT 1320
Db 421 GlnAspLysGlyIleThrIleProHisAspIleAspLeuGlyAspSerArgValIle 440
Qy 1321 CAGGATTATGATAACCAATGAAACAGATCGCGCGAGCTTCTCCAGCCCGCATCGCG 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
Qy 1381 CTTTCTCTGCTCTCGAGCTAATCATGTCTTGGCTCTCTCTCACCGCTCCCGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
Qy 1441 GACCACTCCACTTATGGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 481 AspGlnThrThrTyrGlySerSerThrAsnProMetTyrValSerAspThrValThrLeu 500
Qy 1501 GTTATGTTGCGACCGCGCGCGAGCGCTGCGCGCTGCTCGATTGGACCAAGTCAACA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr 520
Qy 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCTGAAACCTCTTTCTCTCTGCGCG 1620
Db 521 LeuAspGlyArgProLeuThrThrIleGlnTyrSerLysThrPheTyrValLeuPro 540
Qy 1621 CTCGCGGTAGCTCTCTTTCTGGGAGGCGAGCACACATAAGCCGGGTACCTTATAAT 1680

```
Db 541 LeuArgGlyIysLeuSerPheThrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGTAGGACCACTGCTTCGAGATGCGCGGCGGACCGGTGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnIleLeuLeuGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCCACTTACACCACTAGCTGGGTGCTGCTGCCGCTCCATTTCTGCGGTTGCCGTT 1800
Db 581 IleserThrThrThrSerLeuGlyAlaGlyProThrSerIleSerAlaValGlyVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCACTGCTTGCAGATACCTTGCATACCTCCGCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaValLeuGluAspThrValAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGTATGATTTCTGCCGAGAGTCCGCCCTTGGCCCTCAGGCTGCGCTTC 1920
Db 621 HisThrPheAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGAGCTTCAGCGCTTAAGATGAAGTGGGTAAACCTCGGAG 1977
Db 641 GlnSerThrIleAlaGluLeuGlnArgLeuIysMetLysValGlyLysThrArgGlu 659

RESULT 10
US-10-239-090A-53
; Sequence 53, Application US/10239090A
; Publication No. US20040101820A1
; GENERAL INFORMATION:
; APPLICANT: KABUSHIKI KAISHA TOSHIBA
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co
; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same,
; TITLE OF INVENTION: detecting hepatitis E virus using the same
; FILE REFERENCE: 02S0741P
; CURRENT APPLICATION NUMBER: US/10/239,090A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 57
; SEQ ID NO 53
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus JMY-Haw (ORF2)
US-10-239-090A-53

Alignment Scores:
Pred. No.: 1,13e-225 Length: 660
Score: 3174.00 Matches: 607
Percent Similarity: 95.90% Conservative: 25
Best Local Similarity: 92.11% Mismatches: 27
Query Match: 86.98% Indels: 0
DB: 16 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1993) x US-10-239-090A-53 (1-660)
QY 1 ATGCGCCTCGGCCTATTGTTGCTGCTCTCATGTTTTTGGCTATGCTGCGCGGCCA 60
Db 1 MetArgProArgAlaValLeuLeuLeuPheLeuValLeuLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGGTGACCGCTGCGCGCGGTGCTGCGCGCGCGCGCGCGCGGTTCGCGGTGGT 120
Db 21 ProAlaGlyGlnProSerGlyArgArgGlyArgSerGlyGlyThrGlyGly 40
QY 121 TTCTGGGGTGACCGGGTTGATTCACGCTTCGCAATCCCTATATTCATCCCAACCAAC 180
Db 41 PheThrGlyAspArgValAspSerGlnProPheAlaLeuProIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCGATGTCACGCTGCGCGCGCGGCTGACCTGTTGCGCAACCCGCC 240
Db 61 ProPheAlaAlaAspValValSerGlnProGlyAlaGlyAlaArgProArgGlnProPro 80
QY 241 CGACCACTCGGCTCGGCTGCGGTGACACGCGCGCGCGCGCGCGGTGCTACGTCGT 300
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnSerGlnArgProSerAlaAlaProArgArg 100
```

```
QY 301 AGACCTACCAGCTGGGGCGCGCGCTAACCGGGTTCGCTCGGCGCCATGACACCCCG 360
Db 101 ArgSerAlaProAlaGlyAlaAlaProLeuThrAlaValSerProAlaProAspThrAla 120
QY 361 CCAGTGGCTGATGTCGACTCCCGCGGGCCCATCTTGGCGCGCGCAGTATAACTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTTCTTCGTGGCCACCGGCACACTAACCTGGTTCTTTATGCCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGGCTTTTACCCTCTCAGGACGGCACCAATACCATATAAATGGCCACCGAGACTTCT 540
Db 161 AsnProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTGCCGCTGCCACAAATCCGTTTACCCTCGCTGGTCCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValValArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGGGGTTACGGCATCTCCATCTCATCTTCTGGCCACAGACACACCCCGGACG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheThrProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATAAATCCTCAGCGATGTTCTGTATTTTAGTCCAGCCCGGATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGTGGCGTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCTGCTGCTTGTATGCTTTGCATA 840
Db 261 ValGluThrThrGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTCGTAAATTCCTACTACTAATACACCTATACCGGTGCGCTGGGCTGTG 900
Db 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGATTTTCGCAACTTACCCCGGTAAACCAATACCGGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGGTTATTCAGCACTGCTCGCCACCGCTTCTGTCGGGTGCGGAGCGGACTGCGGAG 1020
Db 321 SerArgTyrThrSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACACGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGCT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyThrAsnGly 360
QY 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTTCAACCTTGTGACACTCTGCTT 1140
Db 361 ValGlyGluValGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GGGCGCTCGCGCAGAAATTGATTTCTGCGGTGCGGCGGAGCTGTTCTACTCCCGTCCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCCCAATCGGAGCGGCTGTTAAGTTGTATACATCTGAGAGAATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAATCTCGTGGTTATT 1320
Db 421 GlnAspLysGlyIleThrIleProHisAspIleAspLeuGlyAspSerArgValIle 440
QY 1321 CAGATTATGATAACCAACATGAACAGATCGGCGGACGCTTCTCCAGCCCCCATCGCG 1380
Db 441 GlnAspTyrAspAsnGlnHisGluAsnArgProThrProSerProAlaProSerArg 460
```


QY 721 GCCTCTGAGCTTGATCCCAAGTCAGCGCCTACACTATCGTAACCAAGGCTGGCGCTCC 780
Db
QY 241 AlaserGluLeuValleProSerGluargLeuHisTyrArgAsnGlnGlyTyrPargSer 260
Db
QY 781 CTCGAGACCTCTGGGGTGGCTGAGGAGAGGCTACCTCTGTGCTTGTATGCTTTGCATA 840
Db
QY 261 ValGluThrThrGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeucysile 280
QY 841 CATGGCTCACTCGTAAATTCCTACTAATAACACCCCTATACGGTGCCTCGGGCTGTG 900
Db
QY 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGGCTTACCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGCTC 960
Db
QY 301 AspPheAlaLeuGluLeuGluPheargAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGGTATTCCAGCACTGCTCGCACCGCTTCGCGGTGCGGAGCGGAGTCCCGAG 1020
Db
QY 321 SerArgTyrThrSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCGGCTGCTACCGCTTATGAGGAGCTCTATTTACTAGTACTAATGGT 1080
Db
QY 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyThrAsnGly 360
QY 1081 GTCGGTGAGTCGGCGGGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
Db
QY 361 ValGlyGluValGlyArgGlyleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCGACAGAAATTGATTCGTCGGTGTGCGGAGCTTCTTACTCCGCTCC 1200
Db
QY 381 GlyGlyLeuProThrGluLeuileSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTCTCTCAGCAATGGCGGCGGCTGTTAAGTTGATATACATCTGAGAGAACTCTCAG 1260
Db
QY 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGGTATTGCAATCCCGCATCGCATTCACCTCGGAGAACTCTGCTGTGTTATT 1320
Db
QY 421 GlnAspLysGlyleThrleProHisaspIleAspLeuGlyAspSerArgValValle 440
QY 1321 CAGGATTTATGATAACAACATGAAACAGATCGGCGGAGCTTCTCCAGCGCCATCGGCG 1380
Db
QY 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTCTCTGCTTCGAGCTAATGATGTGCTTTCGCTCTCTCACCGCTGCCGAGTAT 1440
Db
QY 461 ProPheSerValLeuArgAlaAsnAspValleuThrleuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCCACTTATGGCTCTTCGACTGGCGCCAGTTATGTTTCTGACTCTGTGACCTTG 1500
Db
QY 481 ***GlnThrThrTyrGlySerSerThrAsnProMetTyrValSerAspThrValThrLeu 500
QY 1501 GTTAATGTTGCGACGGCGCGAGCGGCTTGGCGGCTGCTCGATTGACCAAGTCTACA 1560
Db
QY 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr 520
QY 1561 CTGAGCGTGGCCCTCTCCACCATCCAGCAGTACTCGAGACCTTCTTGTCTCTGCGCG 1620
Db
QY 521 LeuAspGlyArgProLeuThrThrleGlnGlnTyrSerLysLysPheTyrValleuPro 540
QY 1621 CTCGCGGTAAGCTCTCTTCTGGGAGGCGAGGCAACAATGAGCGGGTACCTTTATAT 1680
Db
QY 541 Leu***GlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACACTGCTAGGACCAACTGCTTGTGAGAAATGCGCGCGGACCGGCTCGCT 1740
Db
QY 561 TyrAsnThrThrAlaSerAspGlnleLeuileGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGCTCCCGTCTCCATTTCTGCGGTTCCGCT 1800
Db
QY 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProThrSer***SerAlaValGlyVal 600

QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGACTACCTCCCGCGCC 1860
Db
QY 601 LeuAlaProHisSerAlaLeuAlaValLeuGluAspThrValAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGTATGATTTCTGCGCAGAGTGGCGCCCTTGGCTTCAGGCTTCGCGCTTC 1920
Db
QY 621 HisThrPheAspAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTGAGCTTACGCGCTTAAAGATGAAGTGGGTAAAACTCGGGAG 1977
Db
QY 641 GlnSerThrIleAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu 659
RESULT 13
US-10-319-745-92
; Sequence 92, Application US/10319745
; Publication No. US20030211467A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushatwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.PI
; CURRENT APPLICATION NUMBER: US/10/319,745
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: US/09/468,147A
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/173,141
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/061,199
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 92
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 481
; OTHER INFORMATION: Xaa = Unknown or Other at position 542
; OTHER INFORMATION: Xaa = Unknown or Other at position 595
US-10-319-745-92
Alignment Scores:
Pred. No.: 1.22e-224 Length: 660
Score: 3160.00 Matches: 604
Percent Similarity: 95.75% Conservative: 27
Best Local Similarity: 91.65% Mismatches: 28
Query Match: 86.60% Indels: 0
DB: 12 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-319-745-92 (1-660)
QY 1 ATGCGCCCTCGCGCTATTTTGTGCTGCTCTCATGTTTTCCTATGTCCTCGCGCA 60
Db
QY 1 MetArgProArgAlaValLeuLeuLeuPheLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCGCGGTCAGCGCTCTGGCGCGCGTCTGGCGCGCGTCTGGCGCGCGTCTGGCGCGGT 120
Db
QY 21 ProAlaGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlyAlaGlyGlyGly 40
QY 121 TTCTGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 180
Db
QY 41 PheTrpSerAspArgValAspSerGlnProPheAlaLeuProTyrIleHisProThrAsn 60
QY 181 CCTTCGCGCCCGGATGTCAACCGCTCGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 240
Db
QY 61 PropheAlaAlaAspValValSerGlnProGlyAlaGlyThrArgProArgGlnProPro 80
QY 241 CGACCACTCGGCTCGCTTGGGTGACCGCGCGCGCGCGCGCTTCACGTCT 300

Db 81 ArgProLeuGlySerAlaThrArgAspGlnSerLysArgProSerValAlaProArgArg 100
QY 301 AGACCTTACCACAGCTGGGGCGCGCGCTAACCGGGTGCCTCGGGCCCATGACACCCCG 360
Db 101 ArgSerThrProAlaGlyAlaAlaProLeuThrAlaIleSerProAlaProAspThrAla 120
QY 361 CCAGTGCCTGATGTCGACTCCCGGGCGCCATCTTGGCGCGCAGTATAACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyAsnLeuSerThr 140
QY 421 TCTCCCTTACTCTTCCGTGGCCACCGGCACCTAACCTGGTCTTTATGCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyAlaAlaProLeu 160
QY 481 AGTCGGCTTTTACCCTTCAGACGGACCAATACCCATATTAATGGCCACGAGCTTCT 540
Db 161 AsnProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTGCCGTGCCCAATCCGTTACCGCCGCTGTCGCCCAAT 600
Db 181 AsnTyAlaGlnTyArgValValArgAlaThrIleArgTyArgProLeuValProAsn 200
QY 601 GCTGTCGGCGGTACGCGCATCTCCATCTCATTTCTGGCCACAGACACACACCCCGCAG 660
Db 201 AlaValGlyGlyTyAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATGATGAATTCATTAACCTCGAGGATGTTCTGATTTAGTCCAGCCGGCATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGTGGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACTCTGGGGTGGCTGAGGAGGAGCTACCTCTGCTTGTATTGATTGGCATA 840
Db 261 ValGluThrThrGlyValAlaGluGluAlaThrSerGlyLeuValIleLeuCysIle 280
QY 841 CATGGCTCACTCGTAAATTCCTATCTACTAATACACCTATACGGTGCCTCGGGCTGTG 900
Db 281 HisGlySerProValAsnSerTyThrAsnThrProTyThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCCTTCGTGCGGTGCGGAGCGGACTCGCGAG 1020
Db 321 SerArgTyThrSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCACAGCGTGTACCGCTTTATGAGGACCTCTATTTACTAGTACTAATGGT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyThrAsnGly 360
QY 1081 GTCGCTGAGTACGCGCGGGATAGCCCTCACCCCTGTTCAACCTTGTCAGACTCTGCTT 1140
Db 361 ValGlyGluValGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTTCGCGCACAGAATTGATTTTCGCGGTGGTGGCCAGCTCTTACTCCCGTCCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTySerArgPro 400
QY 1201 GTTGTCTCAGCAATGGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGATCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrProThrValLysLeuTyThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGATCTCGTGGTATT 1320
Db 421 GlnAspLysGlyIleThrIleProHisAspIleAspLeuGlyAspSerArgValValIle 440
QY 1321 CAGGATTATGATAAACCAATCAACAAGATCGCGCAGCGCTTCTCCAGCCCATCGCGC 1380

Db 441 GlnAspTyAspAsnGlnHisGluAlaAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGCTTCGAGCTAATGATGCTGTTGGCTCTCTCACCGCTCGCGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTy 480
QY 1441 GACCAGTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 481 ***GlnThrThrTyGlySerSerThrAsnProMetTyValSerAspThrValThrLeu 500
QY 1501 GTTAAATGTTCGACGCGCGCCAGCGCTTCCCGGTGCTCGATTGGACCAAGTCAACA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr 520
QY 1561 CTTGACGGTCCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCTCGCG 1620
Db 521 LeuAspGlyArgProLeuThrThrIleGlnGlnTySerLysLysPheTyValLeuPro 540
QY 1621 CTCGCGCTTAAGCTCTCTTCTGGGAGCGACGACCACTAAAGCGGGTACCTTATATAT 1680
Db 541 Leu**GlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyProTyAsn 560
QY 1681 TATAACACCACTGTAGCAGCAACTGCTTCTCGAAGATGCGCGGCGCACCGGTCGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCCACTTACACCACTAGCTGGGTGCTGCTCCGCTCCATTTCTGCGGTTGCCGCT 1800
Db 581 IleSerThrTyThrThrSerLeuGlyAlaGlyProThrSer***SerAlaValGlyVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGTTGAGGATACCTTGGACTACCTGCCCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaValLeuGluAspThrValAspTyProAlaArgAla 620
QY 1861 CATACTTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCTTACGGCTTACGGCTTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTGAGCTTACGGCTTAAGTGAAGTGGTAAACCTCGGAG 1977
Db 641 GlnSerThrIleAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu 659

RESULT 14
US-09-468-147-167
; Sequence 167, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.PI
; CURRENT APPLICATION NUMBER: US/09/468,147A
; EARLIER FILING DATE: 1999-12-21
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 167
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 84
; OTHER INFORMATION: Xaa = Unknown or Other at position 230
; OTHER INFORMATION: Xaa = Unknown or Other at position 447
US-09-468-147-167

Alignment Scores:		5,63e-224	Length:	660
Pred. No.:	3151.00	Matches:	603	
Percent Similarity:	95.45%	Conservative:	26	
Best Local Similarity:	91.50%	Mismatches:	30	
Query Match:	86.35%	Indels:	0	
DB:	10	Gaps:	0	
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-468-147-167 (1-660)				
QY	1	ATGCGCCCTCGCGCTATTTTGTGCTCCTCATGTTTTCCTATGCTGCGCGGCCA	60	
DB	1	MetArgProArgAlaValLeuLeuLeuPheValLeuLeuProMetLeuProAlaPro	20	
QY	61	CGCGCGGTGACCGCTGCTGCGCGCGCTGCTGGCGCGCGAGCGCGGTTCGCGCGTGT	120	
DB	21	ProAlaGlyGlnProSerGlyArgArgGlyArgSerGlyGlyAlaGlyGlyGly	40	
QY	121	TTCTGGGGTGACCGGTTCATTCTCAGCCCTTCGCAATCCCTATATTCATCCCAAC	180	
DB	41	PheTrpGlyAspArgValAspSerGlnProPheAlaLeuProTyrIleHisProThrAsn	60	
QY	181	CCCTTCGCCCGCATGTCACCGCTGCGCGCGGTGCGACCTGCTGCGCAACCCGCC	240	
DB	61	ProPheAlaAlaAspValValSerGlnProGlyAlaGlyThrArgProArgGlnProPro	80	
QY	241	CGACCACTCGCTCGCTGCGGTGACAGGCCCGCGCGCGCGCTGCTGCGTGTGT	300	
DB	81	ArgProLeu**SerAlaTrpArgAspGlnSerGlnArgProSerAlaAlaProArg	100	
QY	301	AGACCTACCACTGCGCGCGCGCTGCGCGCGGTGCGTCCGCGCCCATGACACCCCG	360	
DB	101	ArgSerAlaProAlaGlyAlaAlaProLeuThrAlaValSerProAlaProAspThrAla	120	
QY	361	CGAGTGCCTGATGTCGACTCCGCGCGCGCATTTGCGCGCGCAGTAACTATCAACA	420	
DB	121	ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr	140	
QY	421	TCTCCCTTACCTCTCCGTCGCGCCACCGCCTAACTGCTGCTTTCCTGCGCGCTCT	480	
DB	141	SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu	160	
QY	481	AGTCGCTTTTACCCCTTCAGACCGGCACCAATACCATATATGCGCCACGGAAGTCT	540	
DB	161	AsnProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer	180	
QY	541	AATTATGCGCAGTACCGGTTGCCGTGCCAATCCGTTACGCGCGCTGCTGCCCAAT	600	
DB	181	AsnTyrAlaGlnTyrArgValValArgAlaThrIleArgTyrArgProLeuValProAsn	200	
QY	601	GCTGTCGGCGGTTACGCCATCTCCATCTCTGCGCCACAGCACACACCCCGCGC	660	
DB	201	AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr	220	
QY	661	TCCGTTGATATGAATCAATAACCTCGACGGATGTTTCGTATTTTTCAGTCCAGCCG	720	
DB	221	SerValAspMetAsnSerIleThrSer**AspValArgIleLeuValGlnProGlyIle	240	
QY	721	GCCTGTGACCTGTGATCCCAAGTAGCGCTACACTATCGTAACCAAGCGTGGCGCTC	780	
DB	241	AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer	260	
QY	781	GTTCGAGACCTCTGGGTGCTGAGGAGGAGTCTCTGCTGCTGCTGCTGCTGCTGCT	840	
DB	261	ValGluThrThrGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle	280	
QY	841	CATGGCTCACTGTAATTCCTATACATAACACCTATACCGGTGCGCTGCGGTGTG	900	
DB	281	HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu	300	
QY	901	GACTTGGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGCTC	960	

Db	301	AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal	320
QY	961	TCCCGTTATCCAGCACTCTCTCGCACCGCCTTCGTCGGGTGCGACGGAGTCCGAG	1020
Db	321	SerArgTyrThrSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu	340
QY	1021	CTCACCACACCGCTCTACCGCTTATGAAGGACCTCTATTTTACTAGTACTAATGGT	1080
Db	341	LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheAlaGlyThrAsnGly	360
QY	1081	GTCCGTGAGATCGCGCGCGGATAGCCCTCACCCCTTTCAACCTTCTGACACTGCTT	1140
Db	361	ValGlyGluValGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu	380
QY	1141	GGCGCGCTCGCACAGAATTGATTCGCGGTGGTGGCCAGCTGTTCTACTCCCGTCCC	1200
Db	381	GlyGlyLeuProThrGlnLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro	400
QY	1201	GTTTGTCTCAGCCCAATGGCGGAGCGACTGTTAAAGTTGTATACATCTGTAGAGATGCTC	1260
Db	401	ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln	420
QY	1261	CAGGATTAAGGGTATTGCAATCCGCGATGACATTCGCGAGAACTCGTGTGGTTATT	1320
Db	421	GlnAspLysGlyIleThrIleProHisAspIleAspLeuGlyAspSerArgValValIle	440
QY	1321	CAGGATTATGAACCAACATGAACAAGATCGCGCGCGCTCTCCAGCCCATCGCGC	1380
Db	441	GlnAspTyrAspAsnGln**GluGlnAspArgProThrProSerProAlaProSerArg	460
QY	1381	CTTCTCTCTGCTCTCGAGCTAATGATGCTGCTTGGCTCTCTCACCGCTGCCAGTAT	1440
Db	461	ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr	480
QY	1441	GACAGTCCACTTATGGCTCTTCGACTGGCCCGCAGTTTATGTTCTGACTCTGTGACTTG	1500
Db	481	AspGlnThrThrTyrGlySerSerThrAsnProMetTyrValSerAspThrValThrLeu	500
QY	1501	GTTAATGTTGCCACCGCGCGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1560
Db	501	ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr	520
QY	1561	CTTGCGGTGCGCCCTCTCCACCATCCAGCAGTACTCCAGACCTCTTCTGCTGCGCG	1620
Db	521	LeuAspGlyArgProLeuThrThrIleGlnTyrSerLysThrPheTyrValLeuPro	540
QY	1621	CTCCGGGTGAAGCTCTCTTCTGGGAGGAGCAGCACAACCTAAAGCCGGTACCCCTATAAT	1680
Db	541	LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn	560
QY	1681	TATAACACCACTGCTAGGACCAACTGCTTGTGAGAAATGCGCGCGCGCGCGCTGCT	1740
Db	561	TyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAla	580
QY	1741	ATTTCACCTTACCACTAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1800
Db	581	IleSerThrTyrThrThrSerLeuGlyAlaGlyProThrSerIleSerAlaValGlyVal	600
QY	1801	TTAGCCCGCCACTCTGCGCTAGCTGCTTGGAGTACCTTGGACTACCTGCGCGCGCGC	1860
Db	601	LeuAlaProHisSerAlaLeuAlaValLeuGluAspThrIleAspTyrProAlaArgAla	620
QY	1861	CATACCTTTGATGTTCTTGGCCAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1920
Db	621	HisThrPheAspAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe	640
QY	1921	CAGTCTACTGCTGAGCTTCAGCGCTTAAGATGAAGGTGGGTAAACCTCGCGAG	1977
Db	641	GlnSerThrIleAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu	659

RESULT 15
US-10-319-745-167
; Sequence 167, Application US/10319745


```
Db      521 LeuAspGlyArgProLeuThrThrIleGlnTyrSerLysThrPheTyrValLeuPro 540
QY      1621 CTCGCGGTAGCTCTTTCTGGAGGAGGACAACTAAAGCCGGGTACCTTATAAT 1680
Db      541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY      1681 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGATGCCCGGCGACCGGTCGCT 1740
Db      561 TyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAla 580
QY      1741 ATTTCCACTTACACCACTAGCTGGGTGCTGCTCCCGTCTCCATTTCTGCGGTTCGCGTT 1800
Db      581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProThrSerIleSerAlaValGlyVal 600
QY      1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGGACTACCTGCGCGCGCC 1860
Db      601 LeuAlaProHisSerAlaLeuAlaValLeuGluAspThrIleAspTyrProAlaArgAla 620
QY      1861 CATACTTTTGATGATTTCTGCCAGAGTGCCGCCCTTGGCCTTCAGGGCTGCGCTTC 1920
Db      621 HisThrPheAspAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe 640
QY      1921 CAGTCTACTGTCGCTGAGCTTCAGCGCCTTAAGATGAAGTGGTAAAACTCGGGAG 1977
Db      641 GlnSerThrIleAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu 659
```

Search completed: August 16, 2004, 13:33:17
Job time : 137.324 secs

QY 145 CAGCCCTTCGCAATCCCTATATTCATCCACCA-----ACCCCTTC 186
Db 1157 gAlaProMetProThrProThrLeuValArgProLeuLeuLysLeuValHisSerProSe 1177
QY 187 GCGCCCGGATGTACCGCTCGCGCGGGCTGGACCTCGGTTCGCCAACCGCGCCGACCA 246
Db 1177 rProGluValSerAlaSerAlaProGluAlaProLeuThrIleSerSerProLeuHl 1197
QY 247 CTCGGCTCCGTTGGCGTGACAGGCCCGCGCCCGCTGCTCCTCAGCTGTAGACCT 306
Db 1197 s---ValProSerSerLeuProGlyProAlaSerSerProMetPro-----11 1212
QY 307 ACCACAGCTGGGCGCGCGCTAACCGCGCTG-----CTCCGGCC 348
Db 1212 eProAsnSerSerProLeuAlaSerProValSerSerThrValSerValProLeuSerSe 1232
QY 349 CATGACACCCCGCCAGTGTCTGATGCTCGACTCCCGCGCGCCCATCTTGGCGCGCAGTAT 408
Db 1232 rSerLeuProIleSerValProThrThrLeuProAlaProAlaSer-AlaProLeuThrI 1252
QY 409 AACCTATCAACTCTCCCTTACCTCTTCGGTGGCCACCGGCACCTAACCTGTCTTTAT 468
Db 1252 leProIle-SerAlaProLeuThrValSerAlaSerGlyProAlaLeuLeuThrSerVal 1271
QY 469 GCGCGCCCTCTAGTCCGCTTTTACCCTTCAGGACGGCACCACCAATACCCATATAGGCC 528
Db 1272 ThrProLeuAlaProValProAlaProAlaProGly-Pro-----1285
QY 529 ACGGAAGCTTCTAATTATGCCCCAGTACCGGGTTGCCCGCCACCAATCCGTCGCCCG 588
Db 1286 -----ProSerLeuAlaProSerGI 1292
QY 589 CTGTGCCCAATGCTCTCGCGGTTAG-----CATCTCATCTCATCTCTCG 636
Db 1292 yAlaSerProSerAlaSerAlaLeuThrLeuGlyLeuAlaThrAlaProSerLeuSerSe 1312
QY 637 CCACAGACACCAACC-----ACCCGACGTCGTTGATGATGAATCAAT 680
Db 1312 rSerGlnThrProGlyHiaProLeuLeuAlaProThrSerSerHisValProGlyLe 1332
QY 681 AACCTCGAGGATGTTCTGATTTTAGTCCAGCGCGCATAGCTCTGAGCTTGTGATCCC 740
Db 1332 uAsnSerThr-----ValAlaProAla---CysSerProValLeuValPr 1346
QY 741 AGTGAGCGCCCTAC-----A 755
Db 1346 oAla-SerAlaLeuAlaSerProPheProSerAlaProAsnProAlaProAlaGlnAlaS 1366
QY 756 CTATCGTAACCAAGGCTCGCTCGAGACCTCTGGGGTGGCTGAGGAGGAGCTAC 815
Db 1366 eLeuLeuAlaProAlaSerSerAlaSerGlnAlaLeuAla-----ThrP 1381
QY 816 CTCGTGCTGTGTTGCTTGCATACATGCTCACTCGTAAATTCCTATACTAATAACACC 875
Db 1381 roLeuAlaProMetAlaAlaProGlnThrAla-----IleLeuAlaP 1395
QY 876 CTATACCGGTGCTCGGCTGTGGACTTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTAC 935
Db 1395 roSerProAlaPro-----ProLeu-----AlaProLeuP 1405
QY 936 CCGCGGTAAACACCAATACGCGGTCTCCGTTATTCAGCACTGCTCGCACCGCTTCG 995
Db 1405 roValLeuAlaPro-----SerProGlyAlaAlaProValLeuAlaSerSerGlnT 1422
QY 996 TCGCGGTGCGGAGCTGCGGAGCTCCAGCTCCACACCGCGTG-----1036
Db 1422 hrProValProValMetAlaProSerSerThrProGlyThrSerLeuAlaSerAlaSerP 1442
QY 1037 -----CTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGGTGT 1082
Db 1442 roValProAlaProThrProValLeuAlaProSerSerThrGlnThrMetLeuProAlaP 1462

QY 1083 CGGTGAGATCGCGCGGGGATAGCCCTCACCTGTGTTCAACCTT----- 1125
Db 1462 roValProSerProLeu-----ProSerProAlaSerThrGlnThrLeuAlaLeuAlaP 1480
QY 1126 -----CTGACACTCTGCTGGCGGCTGCGACAGAAATGATTTCTTGGCTGGTG 1177
Db 1480 roAlaLeuAlaProThrLeuGlySerSerProSerGlnThrLeuSerLeuGlyThrG 1500
QY 1178 GCGAGCTGTCTACTCCCGTCCGTTGTCTGACCAATGGCGAGCGCAGCTGTTAAATGTT 1237
Db 1500 lyAsnProGlnGlyProPheProThrGlnThrLeuSerLeuThrProAlaSerSerLeuV 1520
QY 1238 ATACATCTCTAGAGATGCTCAGCAGGATAAGGGTATTGCAATCCCGCATGACATTGACC 1297
Db 1520 alProThrProAlaGlnThr-----LeuSerLeuAlaProGlyProProL 1535
QY 1298 TCGGAGAATCTCGTGTGGTTATTTCAGGATTATGATAACCAACATGAACAGATCGGCCGA 1357
Db 1535 euGlyProThrGlnThrLeu-----SerLeuAlaProA 1546
QY 1358 CGCTTCTCAGCCCATCGCGCTTCTCTGCTCCTCGAGCTAATGATGTCTTTGGC 1417
Db 1546 laProProLeuAlaProAlaSerProValGlyProAlaProAlaHisThr-----L 1563
QY 1418 TCTCTCTCAGCGTCCGAGTATGACAGTCCACTTATGCTCTTCTCGACTGGCCAGTTT 1477
Db 1563 euThrLeuAlaProAla-SerSerSerAlaSerLeuLeuAlaProAlaSerValGlnThr 1582
QY 1478 ATGTTTCTGACTCTGTGACCTTGGTTAATGTTGCGACCGCGCGCGCTTGGCCGTT 1537
Db 1583 LeuThrLeuSer-----ProAlaProVal 1590
QY 1538 CGCTCGATTGGACCAAGTCACTTGACGGTGGCGCCCTCTCCACCATCCAGCAGTACT 1597
Db 1591 ProThrLeuGlyProAlaAlaAlaGlnThrLeuAlaLeuAlaPro-----AlaSerThr 1608
QY 1598 CGAAGACCTTCTTGTCTCGCGTCCGCGTAAAGCTCTTCTCGGAGGCGGACACAA 1657
Db 1609 GlnSerProAlaSerGlnAlaSerSerLeuValValSerAlaSerGly-----AlaAla 1626
QY 1658 CTAAAGCGGGTACCTTATAATTAACACCACTGCTAGCAGCAACCACTGTTGTCGAGA 1717
Db 1627 ProLeuProValThrMetValSerArgLeuPro-----ValSerLys 1640
QY 1718 ATGCCCGCGGCGCGGCTGCTATTTCACCTTACCACTAGCTGGGTGGTGGTCCCG 1777
Db 1641 AspGluProAspThr----- 1645
QY 1778 TCTCCATTTCTCGGTTGCGGTTTAGCCCCCCTCTCGCTAGCATTTGCTTGAGGATA 1837
Db 1646 -----LeuThrLeu-ArgSerGlyProProSerProProSerThrAlaThrSerPh 1662
QY 1838 CTTTGGACTACCTCGCGCCCACTATTTTGTATGATTTCTGCCAGAGTGGCGCCCC 1897
Db 1662 edGly-----ProArgPro-----ArgArgGlnProProPr 1673
QY 1898 TTGGCTTTCAGGCTGCGCTTTCAGTCTAC 1928
Db 1673 oProPro-----ArgSerProPheTyr 1680

RESULT 2

US-10-170-205E-18914
; Sequence 18914, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18914


```
; LENGTH: 1878
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-18914
Alignment Scores:
Pred. No.: 1.44e-05 Length: 1878
Score: 203.00 Matches: 163
Percent Similarity: 37.48% Conservative: 93
Best Local Similarity: 23.87% Mismatches: 269
Query Match: 5.56% Indels: 158
DB: 6 Gaps: 28

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-170-205E-18914 (1-1878)
Qy 13 CCTATTGTTGGTCTCTCATGTTTGGCTATGTCGCCGCGCCACCGCCGGTCAG 72
Db |||:|||||
Dy 224 ProValIleSerAlaAspSerProPheValProValIleProAla-AspGlyProVal-- 242
Qy 73 CCGTCTGGCCCGCTCGTGGCGGCGACGGCGGTTCGGCGGTG---GTTCTGGGGT 129
Db |||:|||||
Dy 243 -----ValValAspSerSerThAspValIleValProValPheProVa 258
Qy 130 GACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAACCCCTTCGCC 189
Db |||:|||||
Dy 258 lAsnGlyIleLeuVal-----ProAlaPheAlaAspGluThrValValPr 273
Qy 190 C-----CGATGTCACCGTGGCGGCGGCGGTGGACCTCGTGTTCGC 231
Db |||:|||||
Dy 273 oLeuAsnProAlaAspSerProValIleProValThrProAspAspLeuValValPr 293
Qy 232 CAACCGCGCCGACCACTCGGCTCGCTGGGTGACACGAGGCCAGCGCCGCGCGTTCGCC 291
Db |||:|||||
Dy 293 oValThrProThrAspCysProValAlaProValThrSerAlaAspCysProValValPr 313
Qy 292 TCAC-----GTGTAGACTACACACAGCTGCGGCGCGCGCGCTAACCC 333
Db |||:|||||
Dy 313 oPheThrProAlaAspAspProIleVal--ProValThrProAlaAspSerProIleVal 332
Qy 334 GCGTTCGCTCGGCGCCATGACACCGCGGAGTGCCTGATGTCAGCTCCGCGGCGCCATC 393
Db |||:|||||
Dy 333 SerValThrProAla--AspSerProValValPro----- 343
Qy 394 TTGCGCGGCGAGTAACTATACATCTCCCTTACCTCTTCCGTGGCCACCGGCACT 453
Db |||:|||||
Dy 344 -----ValThrProAlaAspSerProValValProValThrProAlaAspGly 359
Qy 454 AACCTGGTCTTTATGCGCGCCCTCTTAGTCCGCTTTTACCCCTTCAGGACGGCACCAAT 513
Db |||:|||||
Dy 360 ProValValProValAlaSerAlaAspSerProPheValProValThrProAlaAsp 379
Qy 514 ACCCATATATGCGGCAGGAGCTCTTAATATGCCCAGTACCGGTTGCGCGTCCGACCA 573
Db |||:|||||
Dy 380 ProIleValSerValThrProAlaGluSerProAla--IleProValThrProAlaAsps 399
Qy 574 ATCGTTACCGCGCTGTCGCCAATGCTGTCGGCGGTAGCCCATCTCCATCTCATTC 633
Db |||:|||||
Dy 399 erProValValProValThrProAlaAspSerProValIle-SerValAlaSerAlaAsp 418
Qy 634 TGGCCACAGACACCACCCAGCGGCTCGGTGATATGAATCAATCACTCCGCGGAT 693
Db |||:|||||
Dy 419 SerProValValProValThrProAlaAspThrValValValProValThrProAlaAsp 438
Qy 694 GTTCGTATTTAGTCCAGCCCGGCTAGCCTCTGAGCTTGATCCCAAGTAGCGCCTA 753
Db |||:|||||
Dy 439 GlyProValIlePro-----ValThrProAlaGlu-SerPr 450
Qy 754 CACTATC---GTAAACAGGCTGGCGCTCCGTCGAGACCTCTGGGTGGCTCAGGAGGAG 810
Db |||:|||||
Dy 450 oAlaIleProValThrProAlaAspSerProValValProVal----- 464
Qy 811 GCTACTCTGCTGTTGTTATGCTTTGTGATACATGCTACTCGTAAATTCCTATATACT 870
Db |||:|||||
```

QY 1840 TTGGACTACCTGCCCGCGCCCATCTTTGATGATTTCTGCCAGAGT 1888
| | | | | : : : : :
Db 783 oValThrProAlaAspSerProValLeuSerValThrLeuAlaAspSer 799

RESULT 3

US-60-581-351-7551
; Sequence 7551, Application US/60581351
; GENERAL INFORMATION:
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated With Protein In Plants
; FILE REFERENCE: 38-21 (53372)B
; CURRENT APPLICATION NUMBER: US/60/581,351
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: US 60/479,962
; PRIOR FILING DATE: 2003-06-19
; NUMBER OF SEQ ID NOS: 13980
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7551
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-60-581-351-7551

Alignment Scores:
Pred. No.: 1-52e-05 Length: 763
Score: 201.00 Matches: 163
Percent Similarity: 32.03% Conservative: 58
Best Local Similarity: 23.62% Mismatches: 222
Query Match: 5.51% Indels: 248
DB: 7 Gaps: 37

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-60-581-351-7551 (1-763)

QY 7 CTCGGCCCTATTGTTGTCCTCTCATGTTTTCCTATGTCGCGCGCCACCGCC 66
| | | | | : : : : :
Db 34 ProAlaProProCysTrpMetLeuValSerAlaAlaProProCysProProAlaProPro 53
| | | | | : : : : :
QY 67 GGTGACCGCTCTGGCGCGCTGTCGGCGCGGCGGCGGCTTCGGCGGTTCCTCG 126
| | | | | : : : : :
Db 54 AlaProProLysProLysSerLysAla----- 62
QY 127 GGTGACCGGGTTGATTCTCAGCCCTTC---GCAATCCCTATATTCATCAACCAACCC 183
| | | | | : : : : :
Db 63 -----ProPheProProValProProAlaProProAlaProAlaArgGlu 75
QY 184 TTCGCCCCCGATGTACCGCTGCGCGCGGGCTGACCTGCTGTCGCCAACCGGCCGA 243
| | | | | : : : : :
Db 76 LeuAlaProProLeuProProAlaProProGluAlaProArgGluSerArgProAlaLeu 95
| | | | | : : : : :
QY 244 CCACTCGGCTCGG-----CTTGGCGGTGACGAGCGCCGCGCGCCGCGC 285
| | | | | : : : : :
Db 96 Pro-ProCysProProProProProValValleProAspProProGluProAlaAlaProPr 115
QY 286 GTTGCTACAGTGTAGACTACCAAGCTGGGGCGCGG---CGCTAACCGCGGTGCG 341
| | | | | : : : : :
Db 115 oValProProAlaProAsnSerProProPheProProPheProProAlaProLysPheVa 135
QY 342 TCGGCCCATGACACCCCGCGAGTGCCTGATGTGACTCCCGGGGGCCATCTTGGCGCG 401
| | | | | : : : : :
Db 135 lProAlaProProValProProValProAsn-SerProPheProPro----- 151
QY 402 GCAGTATAACCTATCAATCTCCCTTACCTCTTCGTTGGCCACCGGCACTAACCTGCT 461
| | | | | : : : : :
Db 152 -----PheProProAlaAlaLeuAsnProProAlaProProAla----- 164
QY 462 TCTTTATGCGGCCCTCTTAGTCCGCTTTTACCCCTTACGACGGCAACAATACCATAT 521
| | | | | : : : : :
Db 165 -----ProProLeuAla-----AsnSerProProLeuPro----- 174
QY 522 AATGGCCAGGAAGCTTCTAATTATGCCAGTACCGGGTTGCCCGTGCACCAATCCGTTA 581
| | | | | : : : : :
Db 175 -----ProAlaProProThrProAlaGlyThrProProAla 186

QY 582 CGGCCCGCTGCTCCCAATGCTCTCGCGGTACGCCATCTCCATCTCATTTCTGGCCACA 641
| | | | | : : : : :
Db 186 laAlaProTrpProProValProAlaAlaProLysSerLysProAla--SerProProAla 205
| | | | | : : : : :
QY 642 GACCAACACACCCCGACGCTCGGTTGATATGAATTAACAATCAATCAATCGACGGATTTCTG 701
| | | | | : : : : :
Db 205 xgProProAlaProProMetProAlaThrProMetGluPheProProLeuProProVal- 224
| | | | | : : : : :
QY 702 TTTAGTCCAGCGCGGATAGCTCTGAGCTTGATGCCAAGTGAGCGGCTACACATCG 761
| | | | | : : : : :
Db 225 -----ProProAspProLysSer----- 230
QY 762 TAACCAAGGCTCGGCTCGCTCGAGACCTCTCGGGTGGCTGAGGAGGAGGCTACCTCTGG 821
| | | | | : : : : :
Db 231 -----LysGluThrProProAlaProProAlaPro----- 240
QY 822 TCTTGTATGCTTTGCATACATGCTCACTCGTAATTCCTATACATAATACACCCCTATAC 881
| | | | | : : : : :
Db 241 -----ProLysProProAlaProValP 248
QY 882 CGTGCCCTCGGCTGTTGGACTTTGCCCTTGAGTTGAGTTTCGCAACCTTACCCCGCG 941
| | | | | : : : : :
Db 248 roLeuPro-----ProValProProL 255
QY 942 TAACCAACATACCGGGTCT-----CCCGTTATTCAGCACTGCTCGCCA 986
| | | | | : : : : :
Db 255 euProProValProAsnLyslleProProAlaProProAlaProProValAlaAla 275
QY 987 CGCCCTTCCTCGCGGTGCGGAGCGGACTCGCGAGCTCCGAGCTCACCA-----C 1028
| | | | | : : : : :
Db 275 laValLeuValAlaProCysProProLeuProProLeuProAsnAsnHisProProAlaP 295
QY 1029 CAGGCTGCTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGTTGCTCGGTGA 1088
| | | | | : : : : :
Db 295 ro-ProAlaAlaProValProGlyValProLeu----- 305
QY 1089 GATCGCGCGGATAGCCCTCACCTGTTCAACCTGTTGACACTCTGTTGGCGGCGCT 1148
| | | | | : : : : :
Db 306 ---AlaProLeuProAsnSerHisProProAlaPro-----Pro 317
QY 1149 GCCGACAGATTTGATTTCGTCGGCTGGTGGCGAGCTGTTTCTACTCCCG-----TCCCGT 1202
| | | | | : : : : :
Db 318 SerAlaProValProGlyVal-----ProLeuAlaProLeuProLysSerGlyArg 334
QY 1203 TGTCTCAGCAATGGGAGCCGACTGTTAAGTTGTA-----TACTCTGTAGAGATGC 1256
| | | | | : : : : :
Db 335 ProValSerValTrpLysGlySerPheThrThrLeuSerThrPheCysCysArgValCys 354
QY 1257 TCAGCAGGA-----TAAGGTATTGCAATCCCGATGACATTCACCTCGGAGATCTCG 1310
| | | | | : : : : :
Db 355 SerGlyGluValLeuAlaGlyAlaLeuAsnProSer-----ArgProSerArg----- 370
QY 1311 TGTGTTATTTCAGGATTAATGATCAACCAAGATCGCGAGCCCTTCTCCAGC 1370
| | | | | : : : : :
Db 371 -----SerProLeu-----ThrThrThrProAla-LeuProAlaProL 384
QY 1371 CCCATCG---CGCCCTTCTCTGCTCGAGCTAATGATGCTTTGGCTCTCTCTCAC 1427
| | | | | : : : : :
Db 384 eProProLeuProProLeuProProLeuProLysAsnThrAlaVal-ProProLysPro 404
QY 1428 CGCTCGCAGTATGACCACTCCATTCATGCTCTTCGACTCGCCAGCTTTATGTTTCTGA 1487
| | | | | : : : : :
Db 404 roLeuProProValThrAla----- 410
QY 1488 CTCTGTGACCTTGGTTAATGTTGACCGCGCGGAGCGGCTGCGGTGCTCGATG 1547
| | | | | : : : : :
Db 411 -----LeuAlaPro-----ProLeuPro----- 416
QY 1548 GACCAAGGTACACTTGGAGGTTGCGCCCTCTCT-----CCACCAT 1586
| | | | | : : : : :
Db 417 ---ProLeuAlaProLeuProLysSerProGlyValProProAlaProProLysProProG 436

```
QY 1587 CCAGCAGTACTCGAAGACCTCTTTGTCCTCGCGCTCCGGGTAAAGCTCTCTTTCTG3GA 1646
Db 436 lyIysProTrpThrProProLeuAla----- 445
QY 1647 GGCAGGCACAACTAAAGCCGGGTACCTCTATAATATAACACCACTGTGTAGCGACCAACT 1706
Db 446 --ProAlaProProGluProLysThrValProValLeuProPro-----GlyProSerC 463
QY 1707 GCTTGTGAGAGATCGCGCGGACCGGTGCTATTTCACCTTACACCA----- 1756
Db 463 ys-----ProProSerGluLysProAsnProProAlaProProGluProProG 479
QY 1757 -----CTAGCTGGGTGCTGCTCCGCTCCA-----TTTC 1787
Db 479 luProLysSerSerProAlaLeuProProAlaProProAlaProSerMetProSerAlav 499
QY 1788 TGCAGTTGCGGTTTTAGCCCCCACTCTGCGCTAGCATTTGTTGAGGATACCTTTGGACTA 1847
Db 499 alArgValProProSerProPro-----IlePro----- 508
QY 1848 CCCTGCGCGCGCCATACTTTTGATGATTCTGCCAGAGTGCAGCGCCCTTGGCCTTCA 1907
Db 509 -----ProAlaPro-----ProAlaAlaProA 516
QY 1908 GGGCTGCGCTTTCAGTCTA 1927
Db 516 rgAlaSerMetProAlaLeu 522

RESULT 4
US-10-170-205E-19858
; Sequence 19858, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19858
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-19858

Alignment Scores:
Pred. No.: 2.25e-05 Length: 1006
Score: 199.00 Matches: 199
Percent Similarity: 30.35% Conservative: 55
Best Local Similarity: 23.78% Mismatches: 242
Query Match: 5.45% Indels: 341
DB: 6 Gaps: 44

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-170-205E-19858 (1-1006)
QY 34 ATGTTTTCCTATGCTCCGCGCGCCACCGCGGTGAGCGTCTGGCGCGGTGCTG-- 91
Db 137 MetPheHisThrValSerProGlyProProSerAlaArgProProCysArg-ValProPr 156
QY 92 -----GGCGGCGACGCGCGTTCGCG----- 112
Db 156 oThrThrProLeuAenGlyGlyProGlySerLeuProProGluProProSerValSerGl 176
QY 113 -----GGCGTGGTTTCTGGGGTGACCGGTG-----ATTCTCAGCCCT-- 151
Db 176 nAlaPheProThrLeuAlaGlyProGlyGlyLeuPheProProArgLeuAlaAspProVa 196
QY 152 -----TCGCAATCCCTTATATTCATC-----CAACCAACCCCTTC 186
Db 196 lProSerGlyGlySerSerProArgPheLeuProArgGlyAsnAlaProSerProAl 216
QY 187 GCCCCCGATGTACCG-----CT 204
```

```
Db 216 aProProProProAlaIleSerLeuAsnAlaProSerTyrAsnTrpGlyAlaAlaLe 236
QY 205 GCGCCCGGGGTG-----GACCTCGTGTTCGCCAACCCGCCGCCACCTCGGTGCC 255
Db 236 uArgSerSerLeuValProSerAspLeu-----GlySerProPro-----AlaPr 251
QY 256 G---CTTGGGTGACACGAGCCCGCGCGTTCCTCAGCTCGTAGACCTA----- 307
Db 251 oHisAlaSerSerSerProProSerAspProProLeuPheHisCysSerAspAlaLeuTh 271
QY 308 -----CCACAGCTGGGCGCG----- 322
Db 271 rProProProLeuProProSerAsnAsnLeuProAlaHisProGlyProAlaSerGlnPr 291
QY 323 -----CGCCCTAACCGCGTCTCGTCCGCCCATGACACCCGCCCA 363
Db 291 oProValSerSerAlaThrMetHisLeuProLeuValLeuGlyProLeuGlyGlyAlaPr 311
QY 364 GTGCTCATGTGCACTCCCGCGCGCATCTTGGCGCGGCAGTATA----- 409
Db 311 oThrValGluGlyProGlyAlaProProPheLeuAlaSerSerLeuSerAlaAlaA 331
QY 410 -ACCTATCAACATCTCCCTTACCTCTTCCG----- 439
Db 331 aLysAlaGlnHisProProLeuProProProSerThrLeuGlnGlyArgArgProArgAl 351
QY 440 -----TGCCACCGGCACATAACCTGCTTCTTAT 468
Db 351 agLlnAlaProSerAlaSerHisSerSerSerLeuArgProSerGlnArgArgProArgAr 371
QY 469 GCCGCCCTCTTAGTCCGCTTTTACCCTTCCAGG----- 502
Db 371 gProProThrValPheArgLeuLeuGlyArgGlyProGlnThrProArgArgSerAr 391
QY 503 ----ACGCGCAATAATATGCGCCACGGAAGCTTCTAATATGCCCCAGTACCGG 558
Db 391 gProArgAlaProAlaProValProGlnProPheSerLeu-----ProGluProSe 408
QY 559 GTTCCCGGTGCCACAATCCGTTACCGCGCGTGTGCCCAATGCTGTGCGGGTTTACGCC 618
Db 408 rGlnProIleLeuProSerValLeuSer-----LeuLeuGlyLeuProThrPr 424
QY 619 ATCTCCATCTCATCT-----GCCACACAGACACC 648
Db 424 oGlyProSerHisSerAspGlySerPheAsnLeuLeuGlySerAspAlaHisLeuProPr 444
QY 649 ACCA-----CCCCGACGTCGTTGATATGAATCA 678
Db 444 oProProThrLeuSerSerGlySerProProGlnProArgHisProIle----- 460
QY 679 ATAACCTCGAGGATGTTGCTATTTTAGTCAGCGCGGCATAGCTCTGAGCTTGTGATC 738
Db 461 -----GlnProSerLeuPro----- 465
QY 739 CCAAGTGAGCGCTACACTATCGTAACCAAGCTCGCGCTCCGTCGAGACCTCGGGGTG 798
Db 466 -GlyThrThrSerGlySerLeuSerSerValProGlyAlaProAlaProProAlaAla-- 484
QY 799 GCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGTCATACATAGCTCTGCTGTAAT 858
Db 485 -----SerLysAlaProValValPro-----Se 492
QY 859 TCCTATATAATACACCT----- 877
Db 492 rProValLeuGlnSerProSerGluGlyLeuGlyMetGlyAlaGlyProAlaCysProLe 512
QY 878 -----ATACCGGTGCGCTCG----- 892
Db 512 uProProLeuAlaGlyGlyGluAlaPheProPheProSerProGluGlnGlyLeuAlaLe 532
QY 893 -----GGCTGTTGGACTTTGCCCCCTTGGCTTGGAGTTT 924
```

Db 532 uSerGlyAlaGlyPheProGlyMetLeuGlyAlaLeuProLeuSerLeuGly-- 551
QY 925 CGCAACCTTACCCCGGTACACCAATACGCGGTCTCCGTTATTCCAGCAGTCTGCG 984
Db 552 -----GlnProProSerPro-----LeuLeuAs 560
QY 985 CACCGCTTGTGTCGCGGTGGGACGGAGTCCGAGCTCACCAACA-----CGGTGCT 1038
Db 560 nHisSerLeuPheGlyValLeuThrGlyGlyGlyGlnProProProGluProLeuLe 580
QY 1039 ACCCGCTTTATGAGGACCTCTATTTTACTAGTAAAGTGTGCGTGAGATCGGCCGC 1098
Db 580 uPro-----ProPr 583
QY 1099 GGGATAGCCCTCACCTGTTCACCTTGTGACACTCTGTTGGCGGCTGCCGACAGAA 1158
Db 583 oGlyGlyProGlyPro-----ProLeuAlaProGlyGluProGl 596
QY 1159 TTGATTTCGCGCTGCTGGCCAGCTGTTCTACTCCGCTCCCG-----TTGTCTCA 1209
Db 596 uGlyProSerLeuLeuValAlaSerLeuLeuProProProSerAspLeuLeuProPr 616
QY 1210 GCCAATGGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAACTCTCAGCAGGATAAG 1269
Db 616 oProSerAla----- 619
QY 1270 GGTATTGCAATCCCGCATGACATTTGACCTCGGAGAACTCTGCTGTTATTCCAGGATTAT 1329
Db 620 -----ProProSerAsnLeuLeuAlaSerPheLeuProLe 631
QY 1330 GATAACCAACATGAACAAGATCGCGCAGCCCTTCTCCAGCCCACTCGCGCCCTTCTCT 1389
Db 631 uLeuAlaLeuGlyProThrAlaGlyAspGlyGlyGlySerAlaGluGlyAlaGlyGlyPr 651
QY 1390 GTCTCTCGAGCTAATGATGCTTTGCTCTCTCTCACCGCTG-----CCGAGTATGAC 1443
Db 651 oSerGlyGluProPheSerGlyLeuGlyAspLeuSerProLeuPheProLeuSe 671
QY 1444 CAGTCCA-----CTTATGGCTCTTCAGCTGGCCAGTTATGTTTCTGACTCTGACCC 1497
Db 671 rAlaProProThrLeuLeuAlaLeuAsnSerAla-----LeuLeuAlaAlaThrLeuAspPr 690
QY 1498 TTGGTTAATGTTGCGAGCGCGCGAGCGCTGTCGCGCTCGATGTCGACCAAGGTC 1557
Db 690 O-----ProSerGlyThrPro-----ProGlnPr 698
QY 1558 ACATTGAGGTCGCG-----CCCTCTCACCATCCAGCT-----ACTCG 1599
Db 698 oCysValLeuSerAlaProGlnProGlyProProThrSerSerValThrAlaThrTh 718
QY 1600 AAGACCTTCTTCTCTCGCTCGCTCCGCGGTAAGCTCTCTTTCTGGAGGACGACCAACT 1659
Db 718 rAspProGlyAlaSerSerLeuGlyAlaProSerAsnSerGlyArgProProGlnLe 738
QY 1660 A-----AAGCC 1665
Db 738 uLeuSerProLeuLeuGlyAlaSerLeuLeuGlyAspLeuSerSerLeuThrSerSerPr 758
QY 1666 GGGTACCTCTTAATATAACA-----CCACTGCTAGCGACCAACTGTTGTC 1713
Db 758 oGlyAlaLeuProSerLeuLeuGlnProProGlyProLeuLeuSerGlyGlnLeuGlyLe 778
QY 1714 GAGAAATCGCGCGGACCGGTCGCTATTTCACATT-----ACACCACTAGC 1761
Db 778 uGlnLeuLeuProGlyGlyAlaProProProLeuSerGluAlaSerSerProLeuAl 798
QY 1762 CTGGGTGTCGTCGCTCTCCATTCTGCGGTGTCGTTTACGCCCCCACTCTCGGCTA 1821
Db 798 aCysLeuLeu----- 801
QY 1822 GCATTGCTTGGAGATACCTTGGACTACCTCGCCCGGCCCACTTTTGTATGATTCTGCG 1881
Db 802 -GlnSerLeuGlnIleProProGluGlnProGluAlaProCysLeuProProGluSerPr 821

QY 1882 CCAGAGTGGC--GCCCCCTTGGCCCTTCCAGGCTGCGCTTCCAGTCTA 1927
Db 821 oAlaSerAlaLeuGluProGluProAlaArgProProLeuSerAlaLeu 837

RESULT 5

US-60-581-351-7586
; Sequence 7586, Application US/60581351
; GENERAL INFORMATION:
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated With Protein In Plants
; FILE REFERENCE: 38-21(53372)B
; CURRENT APPLICATION NUMBER: US/60/581,351
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: US 60/479,962
; PRIOR FILING DATE: 2003-06-19
; NUMBER OF SEQ ID NOS: 13980
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7586
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis CDC1551
US-60-581-351-7586

Alignment Scores:

Pred. No.:	3,1e-05	Length:	598
Score:	196.00	Matches:	165
Percent Similarity:	29.55%	Conservative:	43
Best Local Similarity:	23.44%	Mismatches:	222
Query Match:	5.37%	Indels:	274
DB:	7	Gaps:	39

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-60-581-351-7586 (1-598)

QY 43 CTTATGCTGCCCGGCCA-----CGCCCGGTGACGCGTCTGCGCCCGCT 87
Db 10 ProValProAlaProAsnSerProProCysProProAlaProProSer----- 26
QY 88 CTTGGGCGCGCAGCGCGGTTCCGCGGTGTTCTGGGGTGACCGGGTTGATTCTCAG 147
Db 27 -----AlaProSerValProAlaPro 33
QY 148 CCCTTTGCAATCCCTATATTTCATCAACACACCCCTCGCCCGGATCTCACCGTGGG 207
Db 34 Pro-----ValProProAlaProLysSerAsnProPhe-ProProValProProPheAl 51
QY 208 GCGGGGCTGGAGCTCGTGTTCGCCAACCGCGCCGACCACTCGGCTCCGCTGGCGGTGAC 267
Db 51 aPro-----AsnProProAlaProProAlaProProLeuAlaAs 64
QY 268 CAGGCGCCAGCGCCCGCGCTTCGCTCAGCTCGTAGACCTACCA----- 310
Db 64 nSerProProValProProAlaProAlaValProProAlaProIleLysPheTrpGluAr 84
QY 311 -----CAGCTGGGCGCGCGCTTAACCGCGGTGCG----- 340
Db 84 gAlaAlaTrpProProValProAlaAlaProArgAsnLysProAlaSerProProArgPr 104
QY 341 -----CTCGGCGCATGACACCGCGCAGTG 366
Db 104 oProAlaProProValSerArgProAsnProProLeuProProValProProGluProI 124
QY 367 CCTGATGTCAGCTCCCGCGGCGCCATCTTGGCGCGCAGTATACTATCAACATCTCC 426
Db 124 eSerLysAlaAlaProProValPro-----Pr 133
QY 427 CTATCTCTTTCGTGCGCA-----CCGCACTAACCTGGTCTTT 465
Db 133 oValPro---ProTrpProValProMetProAspProPro----- 147
QY 466 TATGCGGCCCTCTTAGTCCGCTTTTACCCCTTCAGGACGGCACAATACCATATAATG 525
Db 148 -ValProProIle-----ProAspArgIleProProAlaPro----- 159

```
Qy 526 GCCACGGAGCTTCTAATTATGATCCAGTACCGGTTGCCGTGCCCAATCCGTTACCGC 585
Db 160 -----ProAspProProSerPro-----ProSerAlaProValAl 171
Qy 586 CCGTGTGTCCCAATGTGTGGGGTTACGCCATCTCCATCTCAATCTGTGGCCACAGACC 645
Db 171 aProTipProProLeu-----ProProLeuProAsnAsnHisProPr 185
Qy 646 ACCACCCCGCGAGCTCGTGTGATGATGAATTAACCTCGAGGATGTCGTATTTTA 705
Db 185 oAlaProProSerAlaProValProGlyVal----- 195
Qy 706 GTCACAGCCGCATAGCTCTGAGCTTGATGCCAAGTAGCGCCTACACTATCGTAAC 765
Db 196 -----ProLeuAlaProLeuProLe----- 202
Qy 766 CAAGCTGGCGCTCCGTCCGACACCTCTGGGTGGCTGAGGAGGCTACCTCTGGTCTT 825
Db 203 ---SerGlyArgProValArg-----AlaTipValGlySerLeuIleAlaLeuArgIl 219
Qy 826 GTTATGC-----TTTGCATACATGGCTCACTCGTAAATTCCTAT 864
Db 219 eCysCysCysArgValCysSerGlyValLeuAlaGlyAlaLeuAsnProSerArgProSe 239
Qy 865 ACTAATACACCTATACCG-----GTGCCCTCGGCTGTGGACTTTGGCCTT 912
Db 239 rSerCysProProLysProProAlaProAlaValProAlaGlyAlaProValProProLe 259
Qy 913 GAGCTTGAGTTTCGCAACTTACCCCGGTA-----ACACCAATACCGGGTCTCC 963
Db 259 u-----ProProLeuProProLeuProLeProLleSerThrProLeuProAlaPr 275
Qy 964 CGTTATTCACACTGTCCGACACCGCTTCGTCCGCTGGGAGCGGACTGCCGAGCTC 1023
Db 275 oProLeuProProLeuProAla-----LeuProThrSe 286
Qy 1024 ACCA-----CCACGGCTGTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAAT 1077
Db 286 iProGlyAlaPro-ProAlaProProValProProAlaProAlaLys----- 301
Qy 1078 GGTGTGCTGATGATCGCGCGGGATAGCCCTCACCTGTGTTCAACCTTGTGCACACTCTG 1137
Db 302 -----AspProProAlaProProAlaProProAlaProLeuSerArgProA 318
Qy 1138 CTTGGCGGCTGCCAGAGAAATGATTTCTGCGCTGTGGCCAGCTCTTCTACTCCCGT 1197
Db 318 laPheProProAlaPro-----ProAla-----Pro- 326
Qy 1198 CCCGTTGTCTCAGCAATGGCGAGCGGACTGTTAAGTTGTATACATCTGTAGAGATGCT 1257
Db 327 Pro-----AlaSerLysProSerProThrProProAlaProProGluProAsnAsnVal 345
Qy 1258 CAGCAGGATAAGGGTATT-----GCAATCCCGCATGACATTTGACCTCGGAGAACTCTCGT 1311
Db 346 ProProSerProProLleProProAlaProProProSerGlyLeu----- 361
Qy 1312 GTGGTTATTCAGGATTATGATACCAACATGACCAAGATCGCGCAGCGCTTCTCCAGCC 1371
Db 362 -----AspProProLeuProProAlaPro 369
Qy 1372 CCATCGCGCCTTTCTCTGTCTCTTC-----GA 1398
Db 370 ProAlaAlaPro-ArgLeuSerMetProAlaSerProProAlaProProPheProProTh 389
Qy 1399 GCTAATGATGTGCTTGGCTCTCTCTCACCGCTG----- 1432
Db 389 rLeuIleMetLeuValProProLeuProProValProProAlaProAsnSerProProGl 409
Qy 1433 ----CCGAGTATGACAGTCCACTTATGGCTCTTCGACTGGCCCGCAGTTTATTTCTGAC 1489
Db 409 uProProSerProProAlaPro----- 416
```

```
Qy 1489 TCTGTGACCTTGGTTAATGTTGACCGCGCGCAGCGCTTGCCTGCGTCTGCTCGATTGG 1548
Db 417 -----ProProLysMetProAsnProProGlyProProValProProAlaProAsnSe 434
Qy 1549 ACCAAGTTCACACTTGACGGTCGCCCTCTCCACCATTC-----AGAGTAC 1596
Db 434 rPro-----ProPheProProAspProProAlaProProValProAlaSerVa 450
Qy 1597 TCGAAGACCTTCTTTGTCTCCCTCGCGGTAAAGTCTCTTCTTTCGGAGGACGACACA 1656
Db 450 lAlaProProAlaProProThrProProSerAlaAsnSerProProPheProProAlaPr 470
Qy 1657 ACTAAGCCGGGTACCCCTTATATATACACCACTGCTAGCAGCACTCTTGTCTGAG 1716
Db 470 oProAlaProPro-----ValAlaProLysAlaAlaAlaAsn----- 482
Qy 1717 AATCCCGCGGACACCGGTCGCTATTTCACATTACCACTAGCTGGTGGTGGTCTCC 1776
Db 483 ---ProProGly-----ProProThrProAlaAla-----Pr 492
Qy 1777 GTCTCCATTTCTGCGGTTGCGGTTTTAGCCCGCCCACTCTCGCTAGCATTCCTTGAGGAT 1836
Db 492 oAsnSerMetProAlaAlaProProAlaProPro----- 503
Qy 1837 ACCTTGAGACTACCTCGCGCGCCCACTACTTTTGTGATGATTTCTGCCAGAGTGCGCCCC 1896
Db 504 -----AlaProProValProValLeuAlaLeuProProAlaProProAlaProPr 520
Qy 1897 CTTG 1900
Db 520 oLeu 521

RESULT 6
US-60-568-073-1197
; Sequence 1197, Application US/60568073
; GENERAL INFORMATION:
; APPLICANT: Huang, Fei
; APPLICANT: Han, Xia
; APPLICANT: Shaw, Peter
; APPLICANT: Clark, Edwin
; TITLE OF INVENTION: COLON CANCER BIOMARKERS
; FILE REFERENCE: 10205 PSP
; CURRENT APPLICATION NUMBER: US/60/568,073
; CURRENT FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 1305
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1197
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-568-073-1197

Alignment Scores:
Pred. No.: 0.000109 Length: 5179
Score: 192.00 Matches: 149
Percent Similarity: 33.69% Conservative: 74
Best Local Similarity: 22.51% Mismatches: 261
Query Match: 5.26% Indels: 178
DB: 33

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-60-568-073-1197 (1-5179)

Qy 143 CTCAGCCCTTCGCAATCCCTATATTCATCAACCA-----ACCCCTTCGCCCCCGC--- 193
Db 1508 IleThrProProAlaSerThrThrLeuProProThrThrThrProSerProProThr 1527
Qy 194 -----ATGCACCGCTGCGCGGGGCTGCACCTCGTG 226
Db 1528 ThrThrThrThrThrProProProThrThrThrProSerProProThrThrThrProle 1547
Qy 227 TTCGCCAACCGCGCAGCACTCGCTCCGCTT---GGCGTGACAGCCCGAGCGCCCG 283
Db 1548 -----ThrProProThrSerThrThrLeuProProThrThrThrProSerProPro 1565
```

```
QY 284 CCGTTGCTCAGCTCGTAGACCTACACAGCTGGGGCGCGCGCTAAACGGCGGTGGCTC 343
Db 1566 ProThrThrThrThrThrProProThrThrThrThrThrProSerProProThrThrThr 1584
QY 344 CGCGCCCATGACACCCCGCAGTGCCTGATGTGCATCCCG-----CGCGCGCATCTTGC 397
Db 1585 ThrProSerProProThrThrThrThrThrThrProProThrThrThrProSerPro 1604
QY 398 GCGGCGAGTATACCATCAACATCTCCCTTACCTTCCGFGGCGACCGGCACTAAC 457
Db 1605 ProThrThrThrThrThrProProProThrThrThrProSerProProThrThrThr 1624
QY 458 TGGTTCTTTATGCGCGCCCTTGTAGTCGGCTTTACCCCTTCAGGACGGCACCAATACC 517
Db 1625 ---ProThrThrProProThrThrThrThrLeuProProThrThrThrProSerPro 1643
QY 518 ATATAATGCGCGGAGCTTCTTAATATGCCAGTACCGGGTTGCCGTGCCACCAATCC 577
Db 1644 Pro-----ProThrThrThrThrProProProThrThrThrProSerProProThr 1661
QY 578 GTTACCGCGGTGTCCTCCCAATGCTGCGGGTTACGGCATCTCCATCTCATTTGCG 637
Db 1662 ThrThrThrProSerProProThrThrThrThrThrThrProProProThrThrThrPro 1681
QY 638 CAC-----AGACCACCACCACCGCGAGCTCGGTGATATGATTCATAA 682
Db 1682 SerSerProThrThrThrProSerProProThrThrThrMet----- 1696
QY 683 CTCGAGGATGTCGTATTTAGTCCACCGCGCATAGCTCTGAGCTTTGTATCCCAA 742
Db 1697 -----ThrThrProSerProThrThrThrProSerProSerPro 1708
QY 743 GTGAGGCTACACTATCGTAACCAAGCTGCGGCTCCGTCGAGACTCTGGGGTGGCTG 802
Db 1709 IleThrThrThrThrProSerSerThrThrThrProSerProPro----- 1724
QY 803 AGGAGGAGCTACCTCTGCTGTGTTATGCTTTGCATACATGCTCACTCGTAAATTCCT 862
Db 1725 -----ProThrThrMetThrThrProSer----- 1732
QY 863 ATACTAATACACCTATACCGGTGCCCTCGGGCTGTGGACTTTGCCCTTGAGCTTGA 922
Db 1733 -----ProThrThrThrProSer-----ProProThrThrThr 1743
QY 923 TTCGCAACCTTACCCCGGTACA-----CCATAGCGGGTCTCCCGTTATTCCA 973
Db 1744 MetThrThrLeuProProThrThrThrSerSerProLeuThrThrThrProLeuProPro 1763
QY 974 GCACTGCTCGCCACCGCTTCTGCTCGCGGTGCGGAGGACTGCGGAGCTCACCAACCG 1033
Db 1764 SerIleThrProProThrPheSerProPheSerThrThrThrProThrThrPro----- 1781
QY 1034 CTGCTACCGGCTTATGAAGGACCTATTTTACTAGTACTAATGGTGTGCGGTGAGATCG 1093
Db 1782 -----CysValProLeu-CysAsnTrpThrGl 1790
QY 1094 GCGCGGGATAGCCCTCACCCGTGTCAACCTTGTGCTGACACTCTGTTGGCGGCTGCCA 1153
Db 1790 yTrpLeuAspSerGly-----LysProAsnPheHisLysPro-GlyGly---Asp 1806
QY 1154 CAGAATTGATTTGTCGGCTGTCGCGCAGCTGTTCTACTCCGCTCCCGTTGTCTCAGCCA 1213
Db 1806 hrGlLeuIleGlyAspValCysGly-----ProGlyTrpAlaAla 1820
QY 1214 ATGGCGCGCCGACTGTTAGTTGTATACATCTGTA----- 1248
Db 1820 snIleSerCysArgAlaThrMetTyProAspValProIleGlyGlnLeuGlyGlnThrV 1840
QY 1249 -----GAGAATGCTCAGCAGGATAAGGTATTG 1276
Db 1840 alValCysAspValSerValGlyLeuIleCysLysAsnGluAspGlnLysProGlyGlyV 1860
```

```
QY 1277 CAATCCCG-----CATGACATTGACCTCGGAGAACTCTCGGTGTGTTA 1318
Db 1860 alIleProMetAlaPheCysLeuAsnTyLeuIleAsnValGlnCysCysGluCysValT 1880
QY 1319 TTCAGGATTATGATACCAACATGACAAAGATCGCGCGAGC----- 1359
Db 1880 hrGln-----ProThrThrMetThrThrThrThr 1890
QY 1360 -----CCTTCTCAGCCCATCGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1411
Db 1890 hrGluAsnProThrProThrThrThrProIleThrThr----- 1304
QY 1412 TTTGGCTCTCTCACCCTCGCGAGTATGACAGTCCACTTATGGCTCTCTGACTGGCC 1471
Db 1905 -----ThrThrValThrProThrProThrProThrGlyThrGlnThrProThrThrP 1923
QY 1472 CAGTTTATGTTTGTACTCTGTACCTTGGTAAAGTTTGCACCGCGCGCGCGCGTGG 1531
Db 1923 rolleThrThrThrThrThrValThrProThrProThrProThrGlyThrGlnThrPro 1943
QY 1532 CCGGTCGCTCGATTGGACCAAGTCACTTGCAGGTGCGCCCTCTCCACCATCCAGC 1591
Db 1943 hrThrThrProThrThr-----ThrThrValThrProThrProThrProThr 1960
QY 1592 AGTACTCGAAGACCTTCTTGTCTGCGCTCGCGGTAAAGCTCTCT----- 1639
Db 1961 GlyThrGlnThrProThrThrThrProIleThrThrThrThrValThrProThrPro 1980
QY 1640 -----TCTGGAGCGAGCACTAAAGCCGGGTACCTTATATATAACACACTG 1693
Db 1981 ThrProThrGlyThrGlnThr-----ProThrThr-----ThrProIle 1993
QY 1694 CTAGCGACCACTGTTGTCGAGATGCGCGCGGACCGGTCGCTATTTCGA---CTT 1750
Db 1994 ThrThrThrThrThrThrValThrProThrProThrProThrGlyThrGlnThrProThr 2013
QY 1751 ACACCACTA-----GCCTGGGTGCTGCTCCGCTCCCACTTCTGGGTTGCGTTTAG 1804
Db 2014 ThrProIleThrThrThrThrValThrProThrPro----- 2026
QY 1805 CCCCCCACTCGCGTAGCATTTGTTGAGATACCTTGGACTACCTGCGCGCGGCCATA 1864
Db 2027 ThrProThr-----GlyThrGlnThrProThrThr-----ProIle 2039
QY 1865 CTTTGTGATGTTCTGCCAGAGTGGCGCCCTTGGCTTCCAGGCTCGGCTTTTCCA 1922
Db 2040 ThrThrThrThrThrVal---Thr-ProThrProThrProThrGlyThrGlnThrPro 2057

RESULT 7
US-60-591-969-185
; Sequence 185, Application US/60591969
; GENERAL INFORMATION:
; APPLICANT: Thiagalingam et al
; TITLE OF INVENTION: Differential Expression of Genes in MSI
; TITLE OF INVENTION: Tumors
; FILE REFERENCE: 1657/2001
; CURRENT APPLICATION NUMBER: US/60/591,969
; CURRENT FILING DATE: 2004-07-28
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 185
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-591-969-185

Alignment Scores:
Pred. No.: 0.000109 Length: 5179
Score: 192.00 Matches: 149
Percent Similarity: 33.69% Conservative: 74
Best Local Similarity: 22.53% Mismatches: 261
Query Match: 5.26% Indels: 178
DB: 33 Gaps: 33
```



```
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 280165
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRY4577_187088C.1.psp
US-10-425-115-280165

Alignment Scores:
Pred. No.:          9,83e-05      Length:          321
Score:             187.50        Matches:         97
Percent Similarity: 37.32%       Conservative:   31
Best Local Similarity: 28.28%    Mismatches:   130
Query Match:       5.14%        Indels:         85
DB:                6           Gaps:          18

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-425-115-280165 (1-321)
QY 77 CTGGCGCGTGTGGCGCGCGCAGCGCGGTTCGGCGGTGTTCTGGGTGACCGGG 136
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 137 TTGATTCTCAGCCCTTCCCAATCCCTATATTCATCCAAACCAACCCCTTCGCCCGCATG 196
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 32 IleAlaProArgProThrProLeuProLeuProGlnProLeuLeuAlaProSerPro 51
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 197 TCACGCTCGCGCGCGGTGACCTGTTGCGCAACCCCGCGCACCACTCGGCTCG 256
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 52 ThrProArgProGlyLeuGlyAlaProThrArgThrProPro-----LeuPro 68
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 257 CTTGGCGTG-----ACAGCGCCAGCGCGCGCGGTTCGCTCAC 295
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 69 LeuProLeuProSerProArgProProArgProProThrProThrProLeuProArg 88
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 296 GTCTAGACCTACACAGCTGGG-----CGCGCGCGTAAACCGCGTCTCGCGG 346
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 89 -----ProThrLeuGlyAlaProSerProAlaArgValProArgProArg 104
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 347 CCATGACACCGCGCGCGGTGATGTCGATCGCGCGCGCATTTGCGCGCGCAT 406
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 105 LeuGlyAlaProArg-----ThrProValLeuProSer-----116
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 407 ATACCTATCAATCTCCCTTACTCTCTCGTGGCGCACCGCATTAACCTGTTCT 466
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 117 -----ProArgValGlyAlaProSerProAlaProProVal-----129
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 467 ATGCGCGCGCTTTAGTCCGCTTTTACCCCTTTCAGGACGGCACCAATACCC---ATATAA 523
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 130 ArgProProGlnLeuGlyAlaProLeuProThrProValAlaProThrProGlnLeuGlu 149
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 524 TGGCCACGAAGCTTCAATTATGCGGAGTACCGGTTGCC---GTGCCAAATCCGTT 580
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 150 AspProSerArgGlyProIleLeuProArgProProLeuLysGlyValProSerThr 169
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 581 ACCGCGCGTGTGCCAATGTGTGCGGTTTACGCCATCTCCA-----TCT 628
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 170 ProProLeuValAlaProLeuGlyAlaProLeuProSerProProValProSerAla 189
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 629 CATTCGGCCACAGACCAACACCC-----CGACGTTCGTTTGATATGAATCAATAA 682
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 190 ArgLysGlyThrProProProProGlnPheArgValProIle-----204
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 683 CTTGACCGATGTTGATATTTAGTCCAGCGCGCATAGCTCTGAGCTTGTGATCCCA 742
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 205 -----ArgSerProValIleProProPro-----212
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 743 GTGAGCGCTACACTATCGTAACCAAGGTGCGCTCGTCCGAGACCTCTGGGGTGGCTG 802
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
QY 213 -----GlnLysGlyIleProProSerAlaProLeu-----222
Db |||:|||||ThrValSerLeuValLeuAlaArgGlyIleIleAlaGlyValValAsp 31
```

```
QY 803 AGGAGGAGGCTACCTCTGCTTGTATGCTTTGCATACATGGCTCACTCGTAAATTCCT 862
Db |||:|||||IleProAlaProArgLeuGly-----ValPro 231
QY 863 ATACTAATACACCTATACCGGTGCGCTGCGGCTGTTGACTTTCCTTGGAGCTTGAGT 922
Db |||:|||||LeuProSerProValGlnGlnProProArgThrProIleLeuProSerArgLys 251
QY 923 TTTCGCA-----ACCTTACCCCGGTAAACCAATACGC---GGGTCTCCCGTTATTC 973
Db |||:|||||GlyAlaProProSerIleProLeuSerProProLysLeuGlyAlaProLeuProSer 271
QY 974 GCACCTGCTCGCCACCGCTTCGTCGCGTGGGAGCGGACTCCCGAGCTCACACACACGG 1033
Db |||:|||||ProLeuLeuProProPro-----ArgLeuGlyAlaProProArgThrProIle 287
QY 1034 CTGCTACCC 1042
Db |||:|||||LeuAlaPro 290

RESULT 9
PCT-US02-22858A-249
; Sequence 249, Application PC/TUS0222858A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: Novel Nucleic Acids and
; SECRETED POLYPEPTIDES
; FILE REFERENCE: 805
; CURRENT APPLICATION NUMBER: PCT/US02/22858A
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: pc.fl_genes version 5.0
; SEQ ID NO 249
; LENGTH: 1250
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-22858A-249

Alignment Scores:
Pred. No.:          0.0002      Length:          1250
Score:             185.50      Matches:         158
Percent Similarity: 29.87%     Conservative:   55
Best Local Similarity: 22.16%  Mismatches:   237
Query Match:       5.08%      Indels:         264
DB:                1          Gaps:          32
```



```
Db 902 GlyGlnGlySerProCysGluGluGlnGluProArgAlaProValAla----- 918
QY 1772 GTCCGCTCTCCATTCTCGCGTTCGCGTTTATAGCCGCCCACTCTG 1816
Db 919 ---ProThr-----ProProThrLeu 924

RESULT 11
US-10-896-891-303
; Sequence 303, Application US/10896891
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/10/896,891
; CURRENT FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: 60/231,328
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 303
; LENGTH: 5262
; TYPE: PRT
; ORGANISM: Human
US-10-896-891-303

Alignment Scores:
Pred. No.: 0.000527 Length: 5262
Score: 182.00 Matches: 138
Percent Similarity: 31.12% Conservative: 54
Best Local Similarity: 22.37% Mismatches: 176
Query Match: 4.99% Indels: 249
DB: 6 Gaps: 31

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-896-891-303 (1-5262)
QY 149 CTTTCGCAATCCCTATATTATCAACAACCCCTTCGCCCGGATGTCACCGCTGGG 208
Db 498 ProGluGluSerProLeu-----SerProProGluSerSerProPheSer 513
QY 209 CCGGGCTGACCTCGTGTTCGCCAACCGCCCGACCACTCGGCTCCGCTT---GGCGTG 265
Db 514 Pro---LeuGluGluSerProLeuSerProProGluGluSerProProSerProAlaLeu 532
QY 266 ACCAGGCCACGCGCCCGCTGTGCTCAGCTCGTAGACACACAGCTGGGCGCGC 325
Db 533 GluThrProLeuSerProPro-----ProGluAlaSerProLeu 545
QY 326 CGCTAACCGCGGTGCTCCGCGCCATGACACCCCGCCAGTGCCTGATGCTGCACTCCCGC 385
Db 546 SerProProPheGluGluSerProLeuSerProProProGluGluLeuProThrSerPro 565
QY 386 GCGCATCTTCGCGCGGAGATACCTATCAACATCTCCCGTTACCTCTCCGTTGGCCA 445
Db 566 ProProGlu---AlaSerArgLeuSerProProProGluGluSerProMetSerPro 584
QY 446 CCGCACTAACTGGTTCTTTATGCGCGCTCTTAGTCCG---TTTACCCCTTCAGG 502
Db 585 ProGluGluSerProMetSerProProProGluAlaSerArgLeuPheProPheGlu 604
QY 503 ACGGCACCAATACCATATAATAGCCACGGAAGCTTCTAATTATGCCAGTACCGGGTTG 562
Db 605 GluSerProLeuSer----- 609
QY 563 CCGTGCCACAATCCGTTACCGCCGCTGTGTCGCCCAATGCTGTGCGCGGTTACGCCATCT 622
Db 610 ProProProGluGluSerProLeuSerProProProGluAlaSerArgLeuSerProPro 629
QY 623 CCATCTCAATCTGGCCACAGACCACACCCAGCTGCGGTGATATGAATCAATAA 682
Db 630 ProGluAspSerProMetSerProProProGluGluSerProMet----- 644
QY 683 CCTCGACGGATGTTGCTATTTTAGTCCAGCCGCGCATAGCCTCTGAGCTTGTGATCCCAA 742
```

```
Db 645 -----SerPro-----ProProGluVal---SerArg 652
QY 743 GTGAGCGCTACACTATCTAACAAGGCTGGCGCTCCGTCGAGA----- 787
Db 653 LeuSerProLeuProValValSerArgLeuSerProProProGluGluSerProLeuSer 672
QY 788 -----CCTCTGGGTGGCTGAGG----- 805
Db 673 ProProAlaLeuSerProLeuGlyGluLeuGluThrProPheGlyAlaLysGlyAspSer 692
QY 806 ---AGGAGGCTACCTCTGCTTGTATGCTTTGCATACATGGCTCACTCGTAAATTCCT 862
Db 693 AspProGluSerProLeuAla----- 701
QY 863 ATACTAATACACCCCTATA---CCGCTGCCCTCGGCTGTTGGACTTTCGCCCTTGAGCTTG 919
Db 702 IleLeuGluThrProIleSerProProGluAlaAsnCysThrAspProGluProVal 721
QY 920 AGTTTCGCAACCTTACCCCGGTAACACAATACGCGGCTCTCCGTTATTCCAGCA--- 976
Db 722 ProProMetIleLeuProPro---SerPro-----GlySerProValGlyProAlaSer 738
QY 977 ---CTGCTCGCCACCGCCTTCGTCGCGGTGCGGACGGAGCTCCGAGCTCACCCACCG 1033
Db 739 ProIleLeu-----MetGluProLeuProProGln 748
QY 1034 CTGCTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGTTCGGTGAGATCG 1093
Db 749 -CysSerProLeuLeuGlnHisSerLeuVal----- 758
QY 1094 GCCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTCTGCTGGCGCCTGCCGA 1153
Db 759 -ProGlnAsnSerPro----- 763
QY 1154 CAGAATTGATTTCGTCGCTGGTGCCAGCTGTTCTACTCCGTCGCCGTTGTCTCAGCCA 1213
Db 764 -----ProSerGlnCys-SerProp 770
QY 1214 ATGGGAGCCGACTGTTAAGTTGTATACATCTGTAGAAATGCTCAGCAGGATAGGTA 1273
Db 770 roAlaLeuPro-----L 774
QY 1274 TTGCAATCCGCATGACATTGACCTCGGAGAATCTCGTGTGTTATTTCAGGATTATGATA 1333
Db 774 euSerValProSerProLeuSerProIleGlyLysValValGlyValSerAspGluAlaG 794
QY 1334 ACCAACATGAACAAGATCGG----- 1353
Db 794 luLeuHisGluMetGluThrGluLysValSerGluProGluCysProAlaLeuGluProS 814
QY 1354 -----CCGACGCTTCTCCAG 1369
Db 814 erAlaThrSerProLeuProSerProMetGlyAspLeuSerCysProAlaProSerProA 834
QY 1370 CCCCATCGCGCCTTCTCTCTCGAGCTAATGATGCTTGTGGCTCTCTCTCACCG 1429
Db 834 laPro----- 835
QY 1430 CTGCGGAGTATGACAGTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTGTGACT 1489
Db 836 --AlaLeuAspAspPheSerGlyLeuGlyGluAspThrAlaProLeu-----AspG 852
QY 1490 CTGTGACCTTGTAAATGTTCCGACCGCGCGGAGCGGTTGCCGCTCGCTCGATTGGA 1549
Db 852 lyIleAsp-----AlaProGlySerGln-ProGluProGlyGlnThrProGly 867
QY 1550 CCAAGGTACACTTGACGGTCGCCCCCTCTCCACCATCCAGCTACTCTGAAGACCTTCT 1609
Db 867 ----- 867
QY 1610 TTGCTCTCGCGCTCGCGGTAAAGCTCTCTTTTCTGGGAGCGGACCAACTAAAGCCGGT 1669
```

```
Db 868 -----SerLeuAlaSerGluLeuLysGlySerProValLeuLeuAspProGlu 883
QY 1670 ACCCTTATATATACACACTG-----CTAGCGACCACTGCTTGTGCGAGATGCCGCCG 1726
Db 884 GluLeuAlaProValThrProMetGluValTyrProGluCys-----LysGlnThrAla 901
QY 1727 GGCACCGCGTGCCTATTT-----CCACTTACACCACTAGCTGGGTGCTG 1771
Db 902 GlyArgGlySerProCysGluGlnGlnGluProArgAlaProValAla----- 918
QY 1772 GTCCCGCTCCATTTCTCGCGTTCGCTTTTAGCCCCCACTCTG 1816
Db 919 ---ProThr-----ProProThrLeu 924

RESULT 12
US-10-851-438-79
; Sequence 79, Application US/10851438
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Hertman, John L.
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-150
; CURRENT APPLICATION NUMBER: US/10/851,438
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/825,751
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/194,314
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 60/225,693
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 79
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Gallus gallus
US-10-851-438-79

Alignment Scores:
Pred. No.: 0.000428 Length: 1151
Score: 180.50 Matches: 163
Percent Similarity: 31.23% Conservative: 65
Best Local Similarity: 22.33% Mismatches: 258
Query Match: 4.95% Indels: 244
DB: 6 Gaps: 33

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-851-438-79 (1-1151)
QY 80 GCCCGCGTGGCGCGCGAGCGGTTTCGGCGGTGCTGCTGGGGTGACCGGGTTG 139
Db 392 AlaAlaThrValProThrAlaGlyAlaValProLysAlaSerThrGlyThr----- 409
QY 140 ATTCTCAGCCCTCGCAATCCCTATATTCATCCACCAACC----- 181
Db 410 -----ProAlaAlaPro-----GlnGlnProValProLysAlaAlaProVal 424
QY 182 -----CCTTCGGCC----- 190
Db 425 ThrProProSerProGlnGlnAlaValProArgAlaAlaThrAlaAlaAlaProVal 444
QY 191 ---CCGATGTCACCGTGGCGCGGGCTGGACCTCGTGTTCGCCAACCCGCCGACCAC 247
Db 445 ThrProGlnGlnProValThrLysAlaAlaThrThrAsnAlaThrProProGln 464
QY 248 TGGCTCCGCTGGCGTGACCGCCAGCGCCCGCGCTGGCTCAGCTGACCTA 307
Db 465 ProfileProLysAlaAlaThrThrThrAlaThrProValThrProGlnGlnProfile 484
```

```
QY 308 CCACAGCTGGGCGCGCGCTAAACGGCGGTGCTCCGGCCCATGACACCCCGCAGTGC 367
Db 485 ProLysAlaGlyThrAspAlaAlaProProAlaValProLysAlaProSerAspGly 504
QY 368 CTGATGTCACCTCCCGCGCGCATCTTCGGCGCGGAGCATATACCTATCAACATCTCCCC 427
Db 505 ArgAlaAlaThrProGlyValProAsnAlaAla-----ThrAspProGlnLysPro 521
QY 428 TTACTCTTCCGTGG-----CCACGGCACTAACTGTTCTTTATGGCGCCCTCTTA 481
Db 522 ProProThrProGlnSerValProSerAlaValThrGluProLysProGlnPro----- 539
QY 482 GTCCGCTTTTACCCCTTCAGGACGCCACCA-----ATA 514
Db 540 -----ArgAlaAlaProProProSerAsnGluAlaThrProAlaVal 553
QY 515 CCCATATAATGGCCACGGAAGCTTCTAATTATGCCAGTACCGGGTTGCCGTG----- 568
Db 554 ProSerProSerProAsnLeuLysSerProLeuProThrIleProLysProValProLeu 573
QY 569 -----CCACAATCCGTTACCGCGCGTGTGCCCAATGCTGTGGGGTTACG 616
Db 574 MetAlaLeuThrProGlnProValThrAlaGlnMetValThrGlnLeuAlaAlaThrLys 593
QY 617 CCATCTCCA-----TCTCATTTCTGGCCACAGACCACCCACCCGACGT 661
Db 594 ProSerProIleValProLysAlaSerProLysAlaLeuMetThrProProProPro 613
QY 662 CCGTTGATATGAATTCATTAACCTCGA-----CGGATGTCGTATTTAG 706
Db 614 ProGlyLeu-----ProArgAlaLeuAlaAlaLysLeuLeuGlyLeuPro 629
QY 707 TCAGGCCG----- 715
Db 630 SerSerProValAlaSerAlaMetHisAlaLysValThrProArgProLeuProAlaSer 649
QY 716 -----GCATAGCTCTGAGCTTGTGATCCCAAGTAGAGCCTACACTATC 760
Db 650 ProValProMetAlaAlaSerProAlaSerLeuGlyProAspAlaAlaArgValAlaLeu 669
QY 761 GTACCAAGGCTGGCGCTCCGTCGAGACCTCTGGGGTGGCTGAGGAGGCTACCTCTG 820
Db 670 AlaThrAsnAlaAlaSerPro-----Gly--AlaLysProGluAlaAlaGlyG 685
QY 821 GTCTTGTATGCTTTCATACATGCTCCTCTCGTA-----AATTCCTACTA 868
Db 685 ly-----AsnGlyThrLeuMetAlaProMetGlyAlaAlaAsnThrG 699
QY 869 ATACACCCCTATACCGGTGCCCTCGGCTGTTGGACTTTCGCCCTTGAGCTTGAGTTTCGCA 928
Db 699 lnMetAlaProIleGlyAlaAlaGlyAlaAlaGlnThrAla----- 712
QY 929 ACCTTACCCCGGTAAACCAATACGGGTCTCCGGTCTTCCTGATTTCCAGACTGTCTGCCACC 988
Db 713 -----ProMetGlyAlaAlaHisThrHisValSer----- 722
QY 989 GCCTTCGTGGCGTGGCGGAGCTGCCGAGCTCACCACTACCGCGGTCTACCCGCTTTA 1048
Db 723 -----ProMetGlyAlaGlyGlyAlaThrGlnMetSerProThrGlyAla----- 737
QY 1049 TGAAGGACCTCTATTTTACTAGTAAATGGTTCGGTGAGATCGCGCGGGATAGCCC 1108
Db 738 -----AlaAsnThrHisMetSerProIleGlyAlaGlyGlyAlaThr 751
QY 1109 TCACCCCTG-----TTCAACCTTGCTGACACTCTGTTGGCGGCTGCCGA 1153
Db 751 hrGlnMetSerProMetGlyAlaAlaAsnThrGlnMetSerProMetGlyAlaThrThr 771
QY 1154 CAGAAATTGATTTTCGTGGCTGGTGGCGCAGCTGTTCTACTCCCGTCCCGTTGCTCAGCCA 1213
Db 771 hrGlnMetSerProMetGlyAlaAlaAlaThrThrGlnProSerProMetGlyAlaAla 791
```

```
QY 1214 ATGGCGAGCGCGCTTAAAGTTGTATACATCTGTAGAGAATGCTCAGCAG-----GATA 1267
Db      :|||
791 laThrGlnValThrAla-----ThrSerAlaGlyAsnThrMetGlnValSerProM 808
QY 1268 AGGGTATTGCAATCCGCGATGACATTGACCTCGGAGAAATCTCGTGGTTATTTCAGGATT 1327
Db      :|||
808 etGlyAlaAlaThrProGln-----
QY 1328 ATGATAACCAACATGAACAAGATCGCGACGCGCTTCT-----CCAG 1369
Db      :|||||
816 -----ThrProSerValGlyAlaAlaThrThrPro 826
QY 1370 CCCATGCGCGCCCTTCTCTGTCCTCGAGCTAATGATGCTTTGGCTCTCTC----- 1423
Db      :|||
826 InProSer-ProMetGlyAlaAlaThrThrLeuMetSerProMetGlyAlaAlaThrThr 845
QY 1424 -----TCACCGCTGCCGAGTATGACCACT-----CCACTTATGGCTCTTCCA 1465
Db      :|||||
846 ProGlnProSerProMetGlyAlaValThrThrGlnProProProMetAlaAlaThrAsn 865
QY 1466 CTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTGTTAAATGTTGCGACCGCGCGCAGG 1525
Db      :|||
866 ThrThrGln-----
QY 1526 CGTTGCGCGGTCGCTCGATTGGACCAAGTCAACATCTGACGCTGCGCCCTCTCCACCA 1585
Db      :|||
869 ProProMetAlaAlaSerThrProGlnSer-----ThrProMetGlyAla 884
QY 1586 TCCAGCAGTACTCGAAGACCTCTTTTGTCTCGCGCTCGCGGTAAAGTCTCTTCTCTGGG 1645
Db      :|||||
885 AlaThrThrThrGlnSerProProMetGlyAlaThrThrThrGlnSerProProMetGly 904
QY 1646 AGGCAGGCACAACTAAAGCGGGTACCCCTT-----ATA 1678
Db      :|||
905 AlaSerThrProGlnAlaProProThrValAlaGlySerProThrProProProProle 924
QY 1679 ATTATACACCACTCTAGCGACCAAC-----TGCCTGTGCGAGAATGCCCGCGGCACC 1732
Db      :|||
925 ProProSerProThrAlaGlnThrSerProGlnProMetSerLysSerProProProAsp 944
QY 1733 GGGTCGCTATTTCOA-----CTTACACCACTAGCTGGGTGCTGTGTCGG 1777
Db      :|||
945 ProProLysAlaProSerAlaAlaGlnThrSerProAlaAlaHisValAlaAlaAsnAla 964
QY 1778 TCTCCATTCTGCGGTGCGGTTTTCAGCCCTTTCAGCCCTTTCAGGAGTTCGTCGAGGATA 1837
Db      :|||||
965 SerProGlyValThr-----
QY 1838 COTTGGACTACCTGCGCGCGCCCATACTTTTGTGATGATTCTGCGCAGAGTGCCTGCGCC 1897
Db      :|||||
970 -----AlaValSerProAlaPro-----IleGlyValThrGluAlaSerPro 983
QY 1898 TTGGCCTTCAGGGCTGCGCTTTCC 1921
Db      :|||
984 SerAlaAspGlyAlaArgLeuSer 991

RESULT 13
US-10-425-115-313976
; Sequence 313976, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 313976
; LENGTH: 708
; TYPE: PRT
```

```
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(708)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_49406C.1.pep
US-10-425-115-313976

Alignment Scores:
Pred. No.: 0.000402 Length: 708
Score: 180.00 Matches: 155
Percent Similarity: 32.72% Conservative: 59
Best Local Similarity: 23.70% Mismatches: 231
Query Match: 4.93% Indels: 210
DB: 6 Gaps: 28

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-425-115-313976 (1-708)
QY 41 TGCCTATGTCGCCGCGCACCGCCCGGTTCAGCCGCTCTGGCGCGCTCGTGGCGCGCGCA 100
Db      :|||
5 CysArgAlaLeuGlnPheHisProAlaAla-----Ser 16
QY 101 GCGCGGTTCCGCGGCTGTTTCTGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCC 160
Db      :|||||
17 ThrProProSerLeuValLys-----LeuSerProProGlnAla 30
QY 161 CCTATTATCATCAACCAACCCCTTCGCCCGCGATGTACCGCTGCGCGCGCGGCTGGAC 220
Db      :|||||
31 ProVal-----GlySerProProProValLysThrThrSerProProAlaPro 47
QY 221 CTCGGTTTCGCAACCGCCGCGACCACTCGGCTCGGCTGGGTGTCACCGAGGCGCGCGCC 280
Db      :|||
48 IleAlaSerProSerProPro-----ProProAlaProValValSerProPro--p 64
QY 281 CGCGCGTTGCTCAGTCGTAGACCTACACAGTGGGGCGCGCGCTAAACCGCGGTG 340
Db      :|||
64 toProValLysSerProProProAlaProValGlySerProProProGluLys 84
QY 341 CTCGGCGCCATGACACCGCGCGAGTGCCTGATGTGCGACTCCCGCGCGCGCATCTTGC 400
Db      :|||||
84 erPro-----ProProAlaPro-ValAlaSerPro----- 94
QY 401 GSCAGTATAACCTATCAACATCTCCCTTACCTCTTCGTTGGCCACCGCSACTAACCTGG 460
Db      :|||
95 -----ProProValLysSerProProProProProProProProProPro 106
QY 461 TTCCTTATGCGCCCTCTTATGTCGCTTTTACCCCTTCAGGACGCGCAACCAATCCCAT 520
Db      :|||
107 ValAlaSerProProProValLysSerProPro-----ProProAlaProVal 123
QY 521 TAATGCGCCAGGAAGCTTCTAATTATGCCAGTACCGGGTTCGCCGTG----- 568
Db      :|||
124 AlaSerProProProValLysSerProProProProProProProProProProProPro 143
QY 569 ---CCACAATCGTTACCGCGCTGGTCCCAATGCTGTGCGCGGTTACGCCATCTCCA 625
Db      :|||
144 ProProAlaProValAlaSerSerProProProMetLysSerProProProProProPro 163
QY 626 TCTCATTTCTGCGCCACAGACCA-----CCACCACCCCGACGTCCTCGTTGATGAATT 676
Db      :|||
164 ValSerSerProProProGluLysSerProProProProProProProProProProPro 183
QY 677 CAATAACTTCGACGGATGTTGTTATTTAGTCCAGCCCGGCATACCTCTGAGCTTGTA 736
Db      :|||
184 ProProProGluGluTy-----ProThrProProProProProProProProPro 196
QY 737 TCCCAAGTGGCGCCTACACTATCGTAACCAAGCTGGCGCTCCGTCGAGACCTCTGGGG 796
Db      :|||
197 -----LysSerSerProProPro----- 202
QY 797 TGGCTGAGGAGGAGGCTACCTCTGCTTGTATTGCTTTTGCATACATGGCTCACTCGTAA 856
Db      :|||
```

Db	203	-----GlulysSerLeuPro	207
Qy	857	ATTCTATATAATA-----CACCTATACCGGTGCCTCGGCTGTGG	901
Db	208	ProThrLeuIleProSerProProGlnGlulysProThrProSerThrPro	227
Qy	902	ACTTTGCCCTTGAGCTTGATTCGCAACTTACCCCGGTAAACACAATACGGGGTCT	961
Db	228	SerLysProProSerProGluLysProSerProLysGluProValSerSerPro	247
Qy	962	CCCGTTATCCAGCACTGCTCGCACCGCTTCGTCGGGTGCGGAGGAGCTCCG--	1018
Db	248	ProGlnThrPro-LysSerSerProProAla--ProValSerSerProProThr	266
Qy	1019	-----AGCTCACCAACCGCTCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTA	1072
Db	267	ProValSerSerProAlaLeuAlaProVal	277
Qy	1073	CTAATGGTGTGGTGCAGATCGGCCGGGATAGCCCTCACCTGTTCAACCTTGCTGACA	1132
Db	278	-----SerSerProProSer	282
Qy	1133	CTCTGCTTGGCGGCTCGCGACAGATTGATTCTGCGGTGGTGGCCAGCTGTCTACT	1192
Db	283	-----ValLysSerSerProPro	288
Qy	1193	CCGTCGCGTGTCTCAGCCAAAT--GGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAG	1251
Db	289	ProAlaProLeuSerSerProProAlaProGlnValLys-----SerSerProPro	306
Qy	1252	AATGCTCAGCAGGATAAGGATTGCAATCCCGCATGACATTGACCTCGGAGAACTCGT	1311
Db	307	ProValGlnValSerSerProProAlaPro-----LysSerSer	320
Qy	1312	GTGGTTATTCAGGANTATGATAACAACTGAACAGATCGCCGACGCTTCTCCAGCC	1371
Db	321	ProProLeuAlaProValSerSerProProGlnValGlulysThrSerProProAla	340
Qy	1372	CCA-----TCGCGCCCTTCTCTGCTCCCTCGAGCT	1401
Db	341	ProLeuSerSerProProLeuAlaProLysSerSerProHisValValSerSer	360
Qy	1402	AATGATGTGCTTTGGCTCTCTCTCACCGCTCCGAGT---ATGACCACTCCACTATAGCC	1458
Db	361	ProProProValValLysSer-SerProProProAlaProValSerSerProProLeuTh	380
Qy	1459	TCITTCAGTGGCCAGTTTATGTTTCTGACTCTGTGACCTTGGTTAATGTTGCGACGGC	1518
Db	380	rProLysProAlaSerProProAlaHisValSerSerProProGluValLysProSe	400
Qy	1519	GCGCAGGCGGTGCGCGTCTCGATTGGACCAAGGTCACTATGACGTCGCGCCCTC	1578
Db	400	rThrProThrProThrThrValLysPro-----ProSe	413
Qy	1579	TCCACCAATCCAGCAGTACTGAAGACCTTCTTTGCTCGCGCTCGCGGTAACTCTCT	1638
Db	413	rGluProLysSerSer-----	418
Qy	1639	TTCTGGGAGGCAGGCACAACTAAAGCCGGGTACCCTTATATTATAACACCACCTGCTAGC	1698
Db	419	-----ProProProThrProValSerLeuProProProLysIleVa	431
Qy	1699	GACCAACTGCTTGTGAGAAATGCCCGCGGCACCGGGTCGCTATTTCCACTTACACCACT	1758
Db	431	lLys-----SerSerProProAlaMetValSerSerProProMetThrProLy	448
Qy	1759	AGCTTGG-----GTGCTGCTCCGCTCTCCATTCTCGGGTTGCGGTTTAAACCCCC	1809
Db	448	sSerSerProProProValValValSerPro-----ProPr	461
Qy	1810	CACCTGCGGTAGCATTCGTTGAGATACCTTGGACTACCTGCCGCGCCCATACTTTT	1869
Db	461	oThrValLys-----SerSerProProAlaPro-----	471

```

QY 1870 GATGATTCTGCCAGAGTGCGCCGCC 1897
Db ::|||:: |||
472 ---ValSerProProAlaThrPro 479

RESULT 14
US-10-425-115-260103
; Sequence 260103, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 260103
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(317)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MKT4577_168832C.l.pap
US-10-425-115-260103

Alignment Scores:
Pred. No.: 0.000372 Length: 317
Score: 179.00 Matches: 85
Percent Similarity: 29.51% Conservative: 18
Best Local Similarity: 24.36% Mismatches: 105
Query Match: 4.91% Indels: 141
DB: 6 Gaps: 18

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-425-115-260103 (1-317)
QY 43 CTATGTGTCGGCGGCACCGCCGGTCAGCGCTGCGCCGCCGTGTCGGCGCGCCAGC 102
Db ::|||:: |||
2 ProMetProProProProProProProPro-----ProAsn 15

QY 103 GCGGTTCGGCGGTGGTTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGAATCCCC 162
Db ::|||:: |||
16 -ProLeuPro-----LeuProAsnPtoLysProAsnPr 26

QY 163 TATATTATCAACAACCCTTCGCCCGCATGTCACGGCTGCGCGCGGGCTGGACCT 222
Db ::|||:: |||
26 oLeuProAsnProProProProProProProProProProProProArgProPr 46

QY 223 CGTGTTCGCCAACCGCCCGCACCTCGCTCGGTTGGCGTGACCGCCAGC----- 277
Db ::|||:: |||
46 oProLeuProAsnProAlaProMetProCysProLeuProProProArgProSerProPr 66

QY 278 -----GCCCGCGGTGGCTTCAGTCTAGACTACCA---CAGCTGGGCGCGCGCG 327
Db ::|||:: |||
66 oProArgProProProLeuProAsnProAlaProMetProCysProLeuProPro----- 84

QY 328 CTAACCGCGGTGCTCCGGCCCATGACACCGCCGACGTCGCTGATGTGACTCCCGCGGC 387
Db ::|||:: |||
85 ---ProArgProProProProProAlaPro***ProLeuProProProArgProProPr 103

QY 388 GCCA-----TCCTGGCGCGGAGTATAACCTATCAACATCTCCC 426
Db ::|||:: |||
103 oProLeuProAsnProProProMetSerCysPro-----IlePr 116

QY 427 CTTAACTTCGTGGCGCACCGGACATAACCTGGTTCTTTATGCGCGCCCTCTTAGTCCG 486
Db ::|||:: |||
116 oProLeuProTrpProProProMetProTrpProMetSerProPro-----132

```



```
QY 487 CTTTACCCTTCAGGACGGACCAATACCCATATAATGGCCACGGAAGCTTCTAATTAT 546
Db 133 -----SerProProLeuProTyPro----- 140
QY 547 GCCAGTACCGGGTTGCCGTCGCCAATCCGTTACCGCCCGCTGTGC-----CCAAT 600
Db 141 -ProProArgProSerProPhePro**ProProLeuMetProTyProSerIleProProPr 160
QY 601 GCTGTGGGGTTACGCCATCTCCATCTCACTTCTGGCCACAGACCACACCCCGGAG 660
Db 160 oLeuPro**ProProProMetPro-----TriProMetProProProSerProPr 177
QY 661 TCCGTTGATAGTAATTCATAACCTCGACGGATGTTGGTATTTAGTCCAGCCCGGCATA 720
Db 177 oProPhe-----Pro**Pr 182
QY 721 GCCTGTGAGCTTGTATCCCAAGTGAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
Db 182 oPro----- 183
QY 781 GTCGAGACTCTGGGGTGG---CTGAGGAGGAGGCTACTCTGGTCTTGTATGCTTTGC 837
Db 184 -----ProMetSerTyProIleProProProLysPro----- 194
QY 838 ATACATGGCTCACTCGTAAATTCCTATATAACACCTATACCGGTGCCCTCGGGCTG 897
Db 195 -----ProProLeuProTyProProProMetPro----- 204
QY 898 TTGGACTTTGGCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAACACCAATACGGG 957
Db 205 -TriProIlePro-----ProProArgProPro----- 213
QY 958 GTCCTCCGTTATCCAGCATCTGCTGCCACCGCTTCGTCGGGGTGGACGGGACTGCC 1017
Db 214 -----ProLeuPr 216
QY 1018 GAGCTCACCAACCGCTGCTACCC 1042
Db 216 o***ProProArgProTriPro 224
RESULT 15
US-10-170-205E-16484
; Sequence 16484, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16484
; LENGTH: 1453
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-16484
Alignment Scores:
Pred. No.: 0.000734 Length: 1453
Score: 177.50 Matches: 185
Percent Similarity: 30.85% Conservative: 68
Best local Similarity: 22.56% Mismatches: 248
Query Match: 4.86% Indels: 319
DB: 6 Gaps: 44
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-170-205E-16484 (1-1453)
QY 37 TTTTGGCTATGCTGCCGCGC----- 57
Db 435 PheLeuProMetAspProAlaThrPheArgArgLysArgProGluSerValGlyGlyLeu 454
QY 58 CCACCGCCCGGT-----CAGCCGCTCTGGCCGCGCTGTCGGCGGCGC 99
```

```
Db 455 GluProProGlyProSerValIleAlaAlaProProSerGly----- 468
QY 100 AGCGCGGTTC----- 114
Db 469 ---GlyGlyAsnIleLeuGlnThrLeuValLeuProProAsnLysGluGlnGluGly 487
QY 115 GGTGGTTTCTGGGGTACCGGGTTGATTCTCAGCCCTTC----- 153
Db 488 Gly-----GlyAlaArgValProSerAlaProAlaProSerLeuAlaTyGlyAla 504
QY 154 ---GCAATCCCTCATATTATCATCCA----- 174
Db 505 ProAlaAlaProLeuSerArgProAlaAlaThrMetValThrAsnValValArgProVal 524
QY 175 -----ACCAACCCCTTC----- 186
Db 525 SerSerThrProValProIleAlaSerLysProPheProThrSerGlyArgAlaGluAla 544
QY 187 GCCCGCGATGTACCGCTCGCGCC-----GGGCTGGACCTCGTCTTCGCCAA 234
Db 545 SerProAsnAspThrAlaGlyAlaArgThrGluMetGlyThrGlySerArgVal---Pro 563
QY 235 CCGCGCCGACCACTCGGCTCGGCTGGCGTGACACGAGCCCGCCCGCTTGCCTCA 294
Db 564 GlyGlySerProLeuGlyValSerLeuValTySerAspLysLysSerAlaAlaThr 583
QY 295 CGTCTAGACCT-----ACCACAGCTGGGGCCCGCGCGCTA 330
Db 584 SerProAlaProHisLeuValAlaGlyProLeuLeuGlyThrValGlyLysAlaProAla 603
QY 331 ACCGCGGTC-----GCTCGGCGCATGACACCCCG----- 360
Db 604 ThrValThrAsnLeuLeuValGlyThrProGlyTyGlyAlaProAlaProProAlaVal 623
QY 360 ----- 360
Db 624 GlnPheIleAlaGlnGlyAlaProGlyGlyGlyThrThrAlaGlySerGlyAlaGlyAla 643
QY 361 -----CCAGTGCCT-----GATCGACTCC 381
Db 644 GlySerGlyProAsnGlyProValProLeuGlyIleLeuGlnProGlyAlaLeuGlyLys 663
QY 382 CGCGGGCCATCTTCGCGCGCAGTATAACCTATCAACATCTCCCTTACCTCTTCCG 441
Db 664 AlaGlyGlyIleThrGlnValGlnPyrIleLeuProThrLeuProGlnGlnLeuGlnVal 683
QY 442 -----GCCACCGGCACATAACCTGGTTCTTTATGCGCGCCCTCTTAGT 483
Db 684 AlaProAlaProAlaProAlaProGlyThrLysAla-----AlaAlaProSerGly 700
QY 484 CCG-----CTTTTACCCCTTCAGGACGGACCAATACCCAT 519
Db 701 ProAlaProThrThrSerIleArgPheThrLeuProProGlyThrSerThrAsnGlyLys 720
QY 520 ATAATGGCCACGGAAGCTTCTAATTATGCCAGTACCGGGTTGCCGTCGCCAATCCGT 579
Db 721 ValLeu-----AlaAlaThr-AlaProThrProGlyIleProIleLeuGlnSerVa 737
QY 580 TACCGCGCGCTGCTCCCAATGCTGCGCGGTATACGCATCTCCATCTCATCTTGGCCA 639
Db 737 lProSerAlaProProLysAlaGlnSerValSerProValGlnAla----- 753
QY 640 CAGACCAACACA-----CCCGACGCTCGGTTGATATGAATATCA 678
Db 754 ---ProProProGlyGlySerAlaGlnLeuLeuProGlyLysValLeuVal----- 769
QY 679 ATAACCTCGAGGATGTTGTTAGTTCAGCCCGGCATAGACCTCTGATGTC 738
Db 770 -----ProLeu----- 771
QY 739 CCAAGTGAGCGCTACACTATGTAACCAAGCTCGCGCTCGTGAGACCTCTCGGGTG 798
```

```
Db 772 -----AlaAlaProSerMetSerValArgGlyGlyAlaGlyGlnPro----- 786
QY 799 GCTGAGGAGAGCTACCTCTGCTGTGTATGCTTTGCATACATGGCTCACTCGTAAAT 858
Db 787 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db -----LeuProLeuValSerProProPheSer-----Va 796
QY 859 TCCT-----ATACTAATACACCTATACCGGT 885
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 796 lProValGlnAsnGlyAlaGlnProProSerLysIlelleGlnLeuThrProValProVa 816
QY 886 G-----CCCTCGGCTGTGGACTTTGCCCTTGAGCTTGAGTTTGCACCTACCCCC 939
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 816 lSerThrProSerGly-----LeuValProProLeuSerProAlaThrLeu--Pro 832
QY 940 GGTAAACACC-----AATACGGGGTCTCCCGTTAT 969
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 833 GlyProThrSerGlnProGlnLysValLeuLeuProSerSerThrArgIleThrTyVal 852
QY 970 TCAGCACTGCTCGCCACCGCTTCGTCCGGT-----GCGGACGGGACTGCCGAG 1020
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 853 GlnSerAlaGlyGlyHisAlaLeuProLeuGlyThrSerProAlaSerSerGlnAlaGly 872
QY 1021 CTCACACACAGGCTGTACCCGCTTTATGAAGACCTCTATTACTAGTACTAATGTT 1080
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 873 ThrValThrSerTyGlyProThrSerSerValAlaLeuGlyPheThrSer----- 889
QY 1081 GTCGGTGAGATCGCGCGGATAGCCTCACCTGTTCAAC----- 1122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 890 LeuGlyProSerGlyProAlaPheValGlnProLeuLeuSerAlaGlyGlnAlaProLeu 909
QY 1123 CTTGCTGCACACTCTGCTGGCGGCTGCCGACA-----GAATTGATT 1164
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 910 LeuAlaProGlyGlnValGlyValSerProValProSerProGlnLeuProAlaCys 929
QY 1165 TCCTCGGCTGTGGCGGCTG-----TTCTAC----- 1191
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 930 AlaAlaProGlyGlyProValIleThrAlaPheTy-SerGlySerProAlaProThrSer 949
QY 1192 TCCCGTCCGCTGTCTCAGCAATGGCGAGCGACTGTTAAGTTGTATACATCTGTAGAG 1251
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 950 SerAlaProLeuAlaGlnProSerGlnAlaProProSerLeuValTyThrValAlaThr 969
QY 1252 AATGCTCAGCAGGATTAAGGTATTGCAATCCCGCATGACATTGACCTCGGAGAACTCGT 1311
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 970 SerThrThrProProAlaAlaThrIleLeuPro----- 980
QY 1312 GTGGTTATTGAGATTATGATACCAACAGATCGCGGAGCGCT----- 1362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 981 -----LysGlyProProAlaProAlaThrAla 989
QY 1363 TCTCCAGCCCATCGCGCCCTTTCTGCTCTTCGAGCTAATGATGTGCTTTGGCTCTCT 1422
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 990 ThrProAlaProThrSerProPheProSerAlaThrAlaGlySerMetThrTy--Ser 1008
QY 1423 CTCACCGCTGCCAGATGACACAGTCCACTTATGGCTTCGACTGGCCCGCTTATGTT 1482
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1009 LeuValAlaProLysAlaGlnArgPro-----SerProLysAlaProGlnLysVal 1025
QY 1483 TCTGACTCTGTGACCTTGTATTATGTTGACCGCGCGGCGGCGCTTGCCTGCTC 1542
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1026 LysAlaAlaIleAlaSerIleProValGlySerPheGluAlaGlyAlaSer-GlyArgPr 1045
QY 1543 GATTGGACCAAGGTCACACTTGACGGTGCCTCTCCACCATCCAGCAGTACTCGAAG 1602
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1045 o---GlyProAlaProArgGlnProLeuGluProGlyPro-----ValArgGl 1060
QY 1603 ACTTCTTTGCTCGCGCTCCGCGTAACTCTCTTTCTGGGAGGAGGCACAACTAAA 1662
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1060 uProThr-----AlaProGluSerGluLeuGluGlyGlnProThrProAl 1076
QY 1663 GCGGGTACCTT-----ATAATTATACACC 1689
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1076 aProProProLeuProGluThrTrpThrProThrAlaArgSerSerProProLeuProPr 1096
```

```
QY 1690 ACTGCTAGCGAACCACTGCTTGTGAGAAATGCGCGGGCACCGGGTCTGCTATTTCACCT 1749
Db 1096 oProAlaGluGluArgThrSerAlaLysGlyProGluThrMetAlaSerLysPheProSe 1116
QY 1750 TACACCACTAGCCTGGGTGCTGCCGTCTCCATTTCTGCGGTTGCCGTTTGTAGCCCC 1809
Db 1116 rSerSerSerAspTrp---ArgValProGlyGlnGlyLeuGluAsnArgGlyGluProPr 1135
QY 1810 CACTCTGCGCTAGCATTTGCTTGAGGATACCTTGAGACTACCTGCGCGGCC 1861
Db 1135 oThr-----ProProSerProAlaProAlaPro 1144
```

Search completed: August 16, 2004, 13:27:00
Job time : 103.28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: August 16, 2004, 13:35:43 ; Search time 89.6821 Seconds
(without alignments)
12476.164 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096
Perfect score: 659
Sequence: 1 ATCGCCCTAGGCTCTTTT.....GTAAACTCGGAGTTGTAG 1980

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues
Word size: 1

Total number of hits satisfying chosen parameters: 2987630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09851410/runat_16082004_125716_9617/app_query.fasta_1.4238
-DB=A Geneseq 29Jan04 -QMT=fastan -SUFFIX=oli.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USFR=US09851410@cgn_1_141/runat_16082004_125716_9617 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A.Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1930s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	659	100.0	659	2	AAR38787 HEV ORF2
2	659	100.0	659	2	AAR38787 HEV ORF2
3	572	86.8	659	2	AAR39308 Mexico st
4	558	84.7	659	2	AAR393387 Human HEV
5	512	77.7	660	2	AAR96090 Hepatitis
6	401	60.8	549	2	AAR96092 Hepatitis
7	392	59.5	540	2	AAR96102 Hepatitis
8	377	57.2	525	2	AAR96104 Hepatitis
9	327	49.6	327	2	AAR96094 Hepatitis
10	327	49.6	327	2	AAR35821 Hepatitis

11	309	46.9	659	5	ABJ04808 Hepatitis
12	149	22.6	549	2	AAW76367 Hepatitis
13	149	22.6	660	2	AAW70323 Hepatitis
14	149	22.6	660	2	AAR70323 Hepatitis
15	149	22.6	660	2	AAR91814 Hepatitis
16	149	22.6	660	2	AAW76369 Hepatitis
17	149	22.6	660	2	AAW93386 Human HEV
18	149	22.6	660	2	AAW93388 Human HEV
19	149	22.6	660	5	AAW93390 Human HEV
20	149	22.6	660	5	ABB81670 Hepatitis
21	149	22.6	660	6	ADA50060 Hepatitis
22	149	22.6	660	7	ADB97807 HEV ORF2
23	146	22.2	660	5	ABJ04820 Hepatitis
24	146	22.2	660	5	ABJ04811 Hepatitis
25	146	22.2	660	5	ABJ04807 Hepatitis
26	146	22.2	660	5	ABJ04809 Hepatitis
27	146	22.2	660	5	ABJ04819 Hepatitis
28	124	18.8	547	2	AAW76369 Hepatitis
29	124	18.8	660	2	AAW93394 Human HEV
30	124	18.8	674	7	ADE06706 Hepatitis
31	124	18.8	674	7	ADE06704 Hepatitis
32	121	18.4	660	5	ABJ04822 Hepatitis
33	121	18.4	660	5	ABJ04821 Hepatitis
34	121	18.4	660	5	ABJ04813 Hepatitis
35	121	18.4	660	5	ABJ04815 Hepatitis
36	120	18.2	525	2	AAR96103 Hepatitis
37	120	18.2	540	2	AAR96101 Hepatitis
38	120	18.2	549	2	AAR96091 Hepatitis
39	120	18.2	660	2	AAR38785 HEV ORF2
40	120	18.2	660	2	AAR39306 Hepatitis
41	120	18.2	660	2	AAR96089 Hepatitis
42	120	18.2	660	2	AAW35826 Hepatitis
43	120	18.2	660	2	AAW80197 Protein e
44	120	18.2	660	2	AAW71210 Human HEV
45	120	18.2	660	2	AAW93389 Human HEV

ALIGNMENTS

RESULT 1
AAR38787
ID AAR38787 standard; protein; 659 AA.
AC AAR38787;
XX
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 11-JAN-1994 (first entry)
XX
DE HEV ORF2 protein.
XX
XX Enterically transmitted non-A non-B hepatitis; ET-NANB; vaccine.
OS Hepatitis E virus; Mexico strain.
XX
XX Key Location/Qualifiers
FT Peptide 225..659
FT Peptide 333..659
FT Peptide /label= SG3
FT Peptide 612..659
FT Peptide /label= 406.3-2
XX
XX WO9314116-A1.
XX
XX 22-JUL-1993.
XX
XX 15-JAN-1993; 93WO-US000459.
XX
XX 17-JAN-1992; 92US-00822335.
PR 01-MAY-1992; 92US-00876941.
XX
XX (GENE-) GENELABS TECHNOLOGIES INC.
PA

```
(USSH ) US SEC DEPT HEALTH.
PA
XX Reyes GR, Bradley DW, Tam AW, Carl M;
XX
XX WPI; 1993-243144/30.
XX DR N-PSDB; AAQ47130.
XX
XX New immunogenic hepatitis E virus (HEV) peptide(s) - are from the ORF1,
XX ORF2 and ORF3 regions of HEV, useful as a vaccine against HEV infection.
XX
XX Disclosure; Fig 8; 48pp; English.
XX
XX Immunogenic hepatitis E virus (HEV) peptides are selected from the ORF1,
XX ORF2 and ORF3 regions of HEV. The peptides can be used in vaccines to
XX prevent infection by HEV. The antibodies can neutralise and block HEV
XX infection and can be used to prevent or treat HEV infection. The peptides
XX and antibodies can also be used as diagnostic reagents. (Updated on 25-
XX MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS
XX field)
XX
XX Sequence 659 AA;
XX
XX Alignment Scores:
XX Pred. No.: 0 Length: 659
XX Score: 659.00 Matches: 659
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR38787 (1-659)
XX
XX 1 ATGCGCCCTAGGCTCTTTGCTGTGTTCCCTGTTCTGCTATGTTGCCGGCCA 60
XX
XX 1 MetArgProArgProLeuLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
XX
XX 61 CGACCGGCTACCGCTCTGCGCGCGCTGCTGGCGCGCGCAGCGGGGTACCGCGGTGT 120
XX
XX 21 ProThrGlyGlnProSerGlyArgArgArgGlyArgSerGlyGlyThrGlyGlyGly 40
XX
XX 121 TTCTGGGTGACGGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC 180
XX
XX 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyraIleHisProThrAsn 60
XX
XX 181 CCCTTTGCCACAGCTTGCCGCTCGTCCGGGTCTGGACCTCGCTTCGCCCAACAGCC 240
XX
XX 61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
XX
XX 241 CGGCCACTTGCTCCACTTGGCGAGATCAGGCCCGCGCCCTCCGCTGCTCCGCTCGC 300
XX
XX 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
XX
XX 301 CGACCTGCCACAGCGGGCTGGCGCTGAGCGGTGGCGCTGGCGCTGCCCATGACACCTCA 360
XX
XX 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
XX
XX 361 CCCGTCGCCGAGTGTGCTTCTCGCGGTGCAATTCTACGCGCGCAGTATAATTGTCTACT 420
XX
XX 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyraAsnLeuSerThr 140
XX
XX 421 TCACCCCTGACATCTCTGTGGCTCTGGCACTAATTTAGTCTCTGTATGACAGCCCTTT 480
XX
XX 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyraAlaAlaProLeu 160
XX
XX 481 AATCGCTCTGCCGCTGACAGCGGTACTAATCTACATTCATATGCGCGACAGAGCGCTCC 540
XX
XX 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
XX
XX 541 AATTATGACAGTACCGGGTCTGCCGCTACTATCCGTTTACCGGCCCTTAGTGGCTAAT 600
XX
XX 181 AsnTyraAlaGlnTyraArgValAlaAlaAlaAlaThrIleArgTyraArgProLeuValProAsn 200
XX
XX 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTCTGGCTCTAAACCAACCAACCCCTACA 660
```

```

XX
XX 201 AlaValGlyGlyTyraAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
XX
XX 661 TCTGTTGACATGAATTCATTCTCCACTGATGTGAGGATTTCTTGTCAACCTGCATCA 720
XX
XX 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
XX
XX 721 GCATCTGAATTCGTCATCCCAAGCGAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 780
XX
XX 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyraArgAsnGlnGlyTrpArgSer 260
XX
XX 781 GTTGAGACATCTCGTGTGTTCTGAGGAGGAAGCCACTCCGGTCTTCTCATGTTATGCATA 840
XX
XX 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
XX
XX 841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCTTGGCTACTCG 900
XX
XX 281 HisGlySerProValAsnSerTyraAsnThrProTyraThrGlyAlaLeuGlyLeuLeu 300
XX
XX 901 GACTTTGCCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAAACACCAATACAGTGTG 960
XX
XX 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
XX
XX 961 TCCCGTTACTCCAGCACTGCTGCTCCTCCGCGCGAGGCGCGAGGAGTCTCGGAGCTG 1020
XX
XX 321 SerArgTyraSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
XX
XX 1021 ACCCAACTGACGACCAAGGTTTCATGAAAGATCTCCACTTACCGGCCCTTAATGGGTTA 1080
XX
XX 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
XX
XX 1081 GGTGAAGTCGCGCGGGATAGTCTTAACATTACTTAACCTTGTGACACGCTCCTCGCGC 1140
XX
XX 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
XX
XX 1141 GGGCTCCCGCACAGAAATTAATTTCTGCTGGCTGGCGGCGCACTGTTTATTTCCGCGCGGT 1200
XX
XX 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyraSerArgProVal 400
XX
XX 1201 GTCTCAGCAATGGCGAGCCCAACCTGAAGCTCTATACATCAGTGGAGATGCTCAGCAG 1260
XX
XX 401 ValSerAlaAsnGlyGluProThrValIysLeuTyraSerValGluAsnAlaGlnGln 420
XX
XX 1261 GATAAGGTTGTGCTATCCCCACCATATCGATCTTGTGTGATTGCTGCTGTGCTCATTCAG 1320
XX
XX 421 AspLysGlyValAlaIleProHisAspLeuLeuGlyAspSerArgValIleGln 440
XX
XX 1321 GATTATGACAAACCCAGCATGAGCAGATCGGCCACCCCGTCCCTCGCCCATCTCGGCT 1380
XX
XX 441 AspTyraAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
XX
XX 1381 TTTTCTGTTCTCCGAGCAATGATGATCTTTGGCTGCTCCCTCCTCCTGACCGCATATGAC 1440
XX
XX 461 PheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGlnTyraAsp 480
XX
XX 1441 CAGTCCACTTACGGTCTGCTCAACTGCGCGGTTTATATCTCGGACAGCGTGTGTTGGTG 1500
XX
XX 481 GlnSerThrTyraGlySerSerThrGlyProValTyraIleSerAspSerValThrLeuVal 500
XX
XX 1501 AATGTTGGACTGGCGCGCAGCGCTAGCCCGATCGCTTGACTGCTGTCACCAAGTCAACCTC 1560
XX
XX 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeu 520
XX
XX 1561 GACGGCGCGCCCTCCCGACTGTTGAGCAATATTCAGAGATTCCTTGTGCTCCCTT 1620
XX
XX 521 AspGlyArgProLeuProThrValGlnGlnTyraSerTyraPhePheValLeuProLeu 540
XX
XX 1621 CGTGCAAGCTCTCCTTTTGGAGCGCGGCAACAAAGAGAGGTATCTCTTATATATAT 1680
XX
XX 541 ArgGlyLysLeuSerPheThrGluAlaGlyThrThrLysAlaGlyTyraProTyraAsnTyra 560
XX
XX 1681 AATCTACTGCTAGTGACGATTCGATTCTGTAATAATGCTGCGCGCATCTCGGCTCGCAT 1740
XX
XX
```



```
QY 901 GACTTTGCTTAGAGCTTTCGATTTTCGCAATCTCACCACTGTATACCAATACACAGTGTG 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTACTCCAGCACTGCTCTCACTCCGCGGAGGGCCGACGGGAGCTGGGAGCTG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCACAATCGACGCCACCACTGTTTCATGAAGATCTCCACTTTTACCGGCTTATATGGGTA 1080
Db 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GGTGAAGTGGCGCGGGATAGCTTAACATTACTTAACCTTGCTCACACGCTCCCTCGGC 1140
Db 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGCTCCGACAGAAATTAATTCGTCGGCTGGCGGCACTGTTTATTCGCCCGCGGTT 1200
Db 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GTCTCAGCCAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGATGCTCAGCAG 1260
Db 401 ValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAGGGTGTGCTATCCCCACGATATCGATCTGGTGATTCGGCTGTGTCATTCAG 1320
Db 421 AspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValValIleGln 440
QY 1321 GATTATGACAACAGCAGCAGGATCGGCCACCCGCTCGCTCGGCATCTCGGCT 1380
Db 441 AspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
QY 1381 TTTTCTGTTCTCCGACAAATGATGTACTTTCGCTGCTCCTCACTCAGCCGAGTATGAC 1440
Db 461 PheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyrAsp 480
QY 1441 CAGTCCACTTACGGTGTCTCACTGGCCCGGTTTATATCTCGACAGCTGATTTGGTG 1500
Db 481 GlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLeuVal 500
QY 1501 AATGTTGGAGCTGGCGCGAGCCGCTAGCCGATCGCTTGAAGTCCCAAGTCAACCTC 1560
Db 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrSerLysValThrLeu 520
QY 1561 GACGGCGGCGCCCTCCGACTGTTGACGAATATTCACAGACATCTTGTGTGCTCCCGCTT 1620
Db 521 AspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuProLeu 540
QY 1621 CGTGCAAGCTCTCTTTTGGAGCGCGGCACAAACAAAGCAGTTATCTTATATAT 1680
Db 541 ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyr 560
QY 1681 AATACTACTGCTAGTGACAGATCTGATTGAAATGCTGCGGCGCATCGGTCGCCATT 1740
Db 561 AsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIle 580
QY 1741 TCAACTATACACAGCTTGGCGCGCTCGGTCGCGCATTTCTCGCGCGCGGTTTGTG 1800
Db 581 SerThrTyrThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeu 600
QY 1801 GCTCCAGCTCGCGCTGCTGCTGAGGATCTTTTGATATTCGCGGCGCGCGGCAC 1860
Db 601 AlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHis 620
QY 1861 ACATTGATGACTTCTGCGCTGAATGCGCGCTTTAGGCCTCCAGGTTGCTCTTCCAG 1920
Db 621 ThrPheAspAppPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGln 640
QY 1921 TCAACTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACCTCGGAGTTG 1977
Db 641 SerThrValAlaGluGlnArgLeuLysValLysValLysTyrArgGluLeu 659
```

RESULT 3

```
AAR39308
ID AAR39308 standard; protein; 659 AA.
XX
AC AAR39308;
XX
DT 25-MAR-2003 (revised)
DT 14-FEB-1994 (first entry)
XX
XX Mexico strain HEV ORF2 putative virus capsid protein.
DE
XX Hepatitis E virus; vaccine; neutralising antibodies; infection; block;
XX open reading frame; antibodies.
OS
XX Hepatitis E virus.
XX
XX WO9314208-A2.
XX
XX 22-JUL-1993.
XX
XX 19-JAN-1993; 93WO-US000475.
XX
XX 17-JAN-1992; 92US-00822335.
XX 20-APR-1992; 92US-00870985.
XX
(GENE-) GENELABS TECHNOLOGIES INC.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Reyes GR, Bradley DW, Twu J, Purdy MA, Tam AW, Krawczynski KZ;
XX
XX WPI; 1993-243223/30.
XX
XX N-PSDB; AAQ46814.
XX
XX Antigen and antibody vaccines against hepatitis E virus infection -
XX contain peptide(s) derived from capsid protein C-terminal or antibodies
XX against protein.
XX
XX Disclosure; Fig 7; 43pp; English.
XX
XX The sequence is that of the putative virus capsid protein encoded by
XX Mexico strain hepatitis E virus (HEV) open reading frame ORF2. This
XX protein or peptide fragments of it may be used in a vaccine composition
XX for immunising an individual against HEV. Antibodies raised against these
XX peptides can also be used in such vaccines. (Updated on 25-MAR-2003 to
XX correct PN field.)
XX
XX Sequence 659 AA;
SQ
Alignment Scores:
Pred. No.: 0 Length: 659
Score: 572.00 Matches: 658
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 86.80% Indels: 2
DB: Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR39308 (1-659)
QY 1 ATGCCCTAGCCCTTTTCTGCTGTTGTTCTCTTCTGCTTCTGCTTCTGCTGCGCA 60
Db 1 MetArgProArgProLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY 61 CCGACCGGTCAAGCCGCTCTGCGCGCGCTGCTGCGCGCGCGCGGTACCGCGGTG 120
Db 21 ProThrGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlyThrGlyGly 40
QY 121 TCTCGGGTGACCGGCTTGTATTCAGCCCTTCGCAATCCCTATATTCATCCAAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTTGCCAGACGTTGCGGCTGCGGCTGCGACCTCGCTTCGCCAACACGAC 240
Db 61 ProPheAlaProAspValAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
```

QY 241 CGGCCACTTGCTCCACTTGGCGAGATCAGCCAGCGCCCTCCGCTGCCCTCCGCTGC 300
 Db |||||
 QY 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
 Db |||||
 QY 301 CGACCTGCCACAGCCGGGGTGGCGGCTGACGGCTGTGGCGCTGCCCATGACACTCA 360
 Db |||||
 QY 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
 Db |||||
 QY 361 CCGCTCCGACAGCTTGATCTCCGGGTGGGAAATCTACGGCGCCAGTATATTTGCTACT 420
 Db |||||
 QY 121 ProValProAspValAspSerArgGlyAlaAlaLeuArgGlnTrpAsnLeuSerThr 140
 Db |||||
 QY 421 TCACCCCTGACATCTCTGCGCTCGGCACATAATTTAGTCTGTATGACGCCGCCCT 480
 Db |||||
 QY 141 SerProLeuThrSerSerValAlaAlaSerGlyThrAsnLeuValLeuThrAlaAlaProLeu 160
 Db |||||
 QY 481 AATCCCGCTCTGCGCTGACGAGCGTACTAATCTACATCTGACATGAGCCAGAGGCTCC 540
 Db |||||
 QY 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
 Db |||||
 QY 541 AATATGACAGTACCGGTTGCCGGCTACTATCCGTTACCGCGCCCTAGTGCCTAAT 600
 Db |||||
 QY 181 AsnTrpAlaGlnTrpArgValAlaAlaArgAlaThrIleArgTrpArgProLeuValProAsn 200
 Db |||||
 QY 601 GCAGTTGGAGCTATGCTATATCAATTTCTTCTGGCTCAAAACAACCAACCCCTACA 660
 Db |||||
 QY 201 AlaValGlyGlyThrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
 Db |||||
 QY 661 TCTGTTGACATGAATTCCTACTTCCACTGATGTCAGGATTCCTTCAACCTGGCATA 720
 Db |||||
 QY 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
 Db |||||
 QY 721 GCATCTGAATGTGATCCCAAGGAGCGCTTCACTACCGCATCAAGTTGGCGCTCG 780
 Db |||||
 QY 241 AlaSerGluLeuValIleProSerGluArgLeuHisTrpArgAsnGlnGlyTrpArgSer 260
 Db |||||
 QY 781 GTTGAGACATCTGTGTGTGAGAGAGAGCCACTCCGCTCTGTGATGTTATGCATA 840
 Db |||||
 QY 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
 Db |||||
 QY 841 CATGGCTCTCCAGTTAACTCCTATACCAATAACCCCTTATACCGGTGCCCTTGGCTACTG 900
 Db |||||
 QY 281 HisGlySerProValAsnSerTrpThrAsnThrProTrpThrGlyAlaLeuGlyLeuLeu 300
 Db |||||
 QY 901 GACTTTCCTTAGAGTTGATTTGCGAATCTACACCTGTATACCAATACACGTGTG 960
 Db |||||
 QY 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrTrpCysAsnThrAsnThrArgVal 320
 Db |||||
 QY 961 TCCGTTTACTCCAGCTGCTGCTCACTCCGCGGAGGGCGGAGCTGCGGAGCTG 1020
 Db |||||
 QY 321 SerArgTrpSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
 Db |||||
 QY 1021 ACCAAGTCCAGCCAGCTTCATGAAGATCTCCACTTTACCGGCTTAATGGGTA 1080
 Db |||||
 QY 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
 Db |||||
 QY 1081 GGTGAAGTCCGCGGATAGCTCTAACATTACTTAACCTGTCTGACACGCTCTCCG 1140
 Db |||||
 QY 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuGly 380
 Db |||||
 QY 1141 GGGCTCCGACAGATTAATTTGTCGCTGCGGGCACTGTTTTATTCCCGCCGGTT 1200
 Db |||||
 QY 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTrpSerArgProVal 400
 Db |||||
 QY 1201 GTCTCAGCAATGGCGAGCCCAACCGTGAAGCTCTATACATCAGTGGAGAATCTCAGAG 1260
 Db |||||
 QY 401 ValSerAlaAsnGlyGluProThrProThrValLysLeuTrpThrSerValGluAsnAlaGln 420
 Db |||||
 QY 1261 GATAAGGTGTGCTATCCCCACATATCGATCTTGATTCGGTGTGCTGCTATTTCAG 1320
 Db |||||
 QY 421 AspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValIleGln 440
 Db |||||
 QY 1321 GATTATGACCAACGATGACAGATCGGCCACCCGCTCGCTCGCCATCTCGGCT 1380
 Db |||||

Db 441 AspTrpAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
 QY 1381 TTTTCTGTTCTCCGAGCAATGATGTACTTGGCTGCTCCCTCACTGCGCCGAGTATGAC 1440
 Db 461 PheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTrpAsp 480
 QY 1441 CAGTCCACTTACCGGTCTCAACTGGCCCGGTTTATATCTCGGACAGCGTACTTTGGTG 1500
 Db 481 GlnSerThrTrpGlySerSerThrGlyProValTrpIleSerAspSerValThrLeuVal 500
 QY 1501 AATGTTCCGACTCGCGCGAGCGCGTAGCCGATCGCTTGACTGTCCTCAAGTTC-ACCCT 1559
 Db 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSer-LysLeuThrLe 520
 QY 1560 CGAGCGGCGCCCTCCGACTCTGAGCAATATCCAGACATTTCTTCTGCTGCCCT 1619
 Db 520 uAspGlyArgProLeuProThrValGluGlnTrpSerLysThrPhePheValLeuProLe 540
 QY 1620 TCGTGGCAAGCTCTCTCTTTGGGAGCGCGGCACAAACAAAGCAGGTTATCTTATATTA 1679
 Db 540 uArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTrpProTrpAsnTy 560
 QY 1680 TAATACTACTGCTAGTGACCCAGATTCTGATGTAATAATGCTCGCGGCCATCGGTGCGCAT 1739
 Db 560 rAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIle 580
 QY 1740 TTCAACCTATACACAGGCTTGGGCGCTCGGTCGCCATTTCTCGCGCGCGGTTT 1799
 Db 580 eSerThrTrpThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLe 600
 QY 1800 GGCTCCACGCTCCGCGCTGCTCTGCTGAGAGATCTTTTGTATTATCCGGCGCGGCGCA 1859
 Db 600 uAlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTrpProGlyArgAlaHi 620
 QY 1860 CACATTTGATGACTTCTGCCCTGAATCCGCGCTTTAGGCTCCAGGTTGCTTTCCA 1919
 Db 620 sThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGl 640
 QY 1920 GTCACATGCTCGTGAGCTCCAGCGCTTAAAGTTAAGTGGTAAACTCGGAGTTG 1977
 Db 640 nSerThrValAlaGluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 659
 RESULT 4
 AAW93387
 ID AAW93387 standard; protein; 659 AA.
 XX AAW93387;
 AC
 XX 11-JUN-1999 (first entry)
 DT
 XX
 DE Human HEV ORF 2 protein from strain Mexico.
 XX
 KW Swine hepatitis E virus; HEV; cross-reaction; antibody; human; therapy;
 XX vaccine; immunise; infection; detection; diagnosis; prevention.
 XX
 OS Hepatitis E virus.
 XX
 PN WO9904029-A2.
 XX
 PD 28-JAN-1999.
 PF 17-JUL-1998; 98WO-US014665.
 XX
 PR 18-JUL-1997; 97US-0053069P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Meng X, Emerson SU, Purcell RH;
 XX
 DR WPI; 1999-132270/11.
 XX
 PT New isolated swine hepatitis E virus - used to develop products for the

PT diagnosis, prevention and treatment of hepatitis E virus infection in
XX mammals, particularly humans.

Example 1; Fig 3A; 70pp; English.

CC This invention describes a swine hepatitis E virus (HEV) and its natural
CC mutants which are capable of cross-reacting with antibodies reactive with
CC a human HEV strain or natural mutants. The HEV and the proteins can be
CC used in vaccines for immunising against HEV infection. The swine HEV can
CC be used in humans to prevent possible infection by human HEV. The swine
CC HEV can also be used as a therapeutic treatment for infection by other
CC strains of HEV. The swine HEV can also be used for the production of
CC antibodies which can be used in therapy, detection and diagnosis. The
CC products can also be used for determining the susceptibility of cells or
CC organs to infection with swine HEV. The swine HEV is particularly useful
CC for the development of agents for the prevention, treatment and detection
CC of human HEV because it is not a human virus and thus can be handled both
CC experimentally and clinically without fear of severe infection and/or
CC contamination

SQ Sequence 659 AA;

Alignment Scores:

Pred. No.: 0 Length: 659
Score: 558.00 Matches: 658
Percent Similarity: 99.85% Conservative: 0
Best Local Similarity: 99.85% Mismatches: 1
Query Match: 84.67% Indels: 0
DB: 2 Gaps: 0

US-09-851-410a-10_COPY_5117_7096 (1-1980) x AAW93387 (1-659)

QY 1 ATGCGCCCTAGGCGCTCTTTGCTGTTGTTCTCTGCTTGTCTGCTATGTTGCCGCGCA 60
DB 1 MetArgProArgProLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY 61 CGACCGGTACCGTCTGCGCGCGTCTGCGCGCGCGAGCGCGGTACCGCGGTGT 120
DB 21 ProThrGlyGlnProSerGlyArgArgArgGlyArgSerGlyGlyThrGlyGly 40
QY 121 TTCTGGGTGACCGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCTTTGCCCCAGACGTTGCCGCTGCTCGCGGTCTGGACCTTCGCGCAACAGCC 240
DB 61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY 241 CGGCGACTTGCTCCACTGGCGAGATCAGGCGCGCGCGCTCGCTGCTCGCGTGC 300
DB 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArg 100
QY 301 CGACCTGCCACAGCGCGCTGCGCGCTGACGGCTGTGGCGCTTCGCGCAACAGCC 360
DB 101 ArgProThrAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 120
QY 361 CCGCTCCCGAGCGTTGATTCTCGCGGTGCAATTCACCGCGCGAGTAAATTTGTCTACT 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCACCCCTGACATCTCTGCGCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 480
DB 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AATCGCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 540
DB 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATATGACAGTACCGGTGCGCGCTACTATCCGTTACCGCGCTTACCGCGCTTAC 600
DB 181 AsnTyrAlaGlnTyrArgValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 200
QY 601 GCAGTTGAGGCTATGCTATATCCATTCTTCTGCGCTCAACACCAACCCCTTACA 660

DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCTGTTGACATGAATCCATTACTTCCACTGATGTCAGGATCTCTTCAACCTGCATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCATCTGAATTCGTCTATCCCAAGCAGCCCTTCACTACCGCAATCAAGCTGGCGCTCG 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTTGAGACATCTGTTGCTCTGAGGAGGAGCCACCTCCGCTCTTCTCATGTATGCATA 840
DB 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCTTGGCTACTG 900
DB 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACCACTGTAAACACCAATACACGTGTG 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCGTTACTCCAGCACTGCTCTCATCTCCGCCCGAGGGCCGACCGGAGCTCGGAGCTG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCCACTGCGAGCCAGCTTCATGAAGATCTCCACTTTACCGGCCCTTAATGGGTA 1080
DB 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GGTGAAGTCGCGCGCGGATAGTCTTAACATTACTTAACCTTGTGACACGCTCTCTCGC 1140
DB 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGCTCCCGACAGAATAATTTCGCTGCGCTGCGCGCAACTGTTTATTCGCGCGGTT 1200
DB 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GCTCAGCAATGCGGAGCCACCGTGAAGCTCTATATATCATCATGTCGAGATCTCAGCAG 1260
DB 401 ValSerAlaAsnGlyGluProThrValIysLeuTyrThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAGGCTGTTGCTATCCCGATATCCCCACGATATCGATCTTGTGTGATTCGCTGCTCATTCAG 1320
DB 421 AspLysGlyValAlaIleProHisAspLeuLeuGlyAspSerArgValIleGln 440
QY 1321 GATTATGACACCGAGCATGAGGATCGGCCACCGCGCTGCGCTGCGCTGCGCTGCGCT 1380
DB 441 AspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
QY 1381 TTTTCTGTTCTCCGAGCAATGATGCTATTTGGCTGCTCCCTCACTGCGAGCGGATGAC 1440
DB 461 PheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGlnTyrAsp 480
QY 1441 CAGTCCACTTACGGCTCGTCAACTGGCGCGCTTATATCTCGGACAGCGTGTGCTGTTG 1500
DB 481 GlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLeuVal 500
QY 1501 AATGTTGCGACTGCGCGCGCGCGCTGACCGCGATGCTGCTGCTGCTGCTGCTGCTGCT 1560
DB 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeu 520
QY 1561 GACGGCGCGCTCCCGACTGTTGAGCAATATTCACAGACATCTTGTGCTGCCCTT 1620
DB 521 AspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuProLeu 540
QY 1621 CGTGGCAAGCTCTCTTTTGGAGCGCGGACCAACAAAGCAGGTATCTCTTATAATAT 1680
DB 541 ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyr 560
QY 1681 AATCTACTGCTAGTGACGAGATCTGATTGAAATGCTGCGCGCTCAACACCAACCCCTTACA 1740

Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100
Qy 634 TGGCTCAACACCAACACCCCTACATCTGTGTGACATGAATCCATTACTTCCACTGAT 693
Db 101 TrpProGlnThrThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp 120
Qy 694 GTGAGGATCTTGTTCAACCTGGCATAGCATCTGAATGTGCATCCCAAGCGAGCGCTT 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGlnLeuValIleProSerGluArgLeu 140
Qy 754 CACTACCGCAATCAAGGTGGCGTTCGCTGAGACATCTGGTGTGTGTGAGGAGGAAGCC 813
Db 141 HisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGluGluAla 160
Qy 814 ACCTCCGGTCTTGTCATGTTATGATCATATGATGCTCTCAGTTAACTCTATACCAATACC 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThr 180
Qy 874 CCTTATACCGGTGGCTTACTGACCTTTCGCTTAGAGCTTGAGTTTCGCAATCTC 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuPheArgAsnLeu 200
Qy 934 ACCACCTGTAAACACCAATACACGTGTGTCCTGTTACTCCAGCACTGCTCGTCA----CTC 989
Db 201 ThrThrCysAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
Qy 990 CGCCGAGGGCGGACCGGAGCTCGGAGCTGACACAACTGACGACCGACGAGTTTCATGAA 1049
Db 221 Arg-ArgGlyAlaAspGlyThrAlaGluLeuThrThrThrAlaAlaThrArgPheMetLy 240
Qy 1050 AGATCTCCACTTACCGGCTTAATGGGTAGGTGAAGTCGCGCGCGGATAGTCTAAC 1109
Db 240 sAspLeuHisPheThrGlyLeuAsnGlyValGlyValGlyArgGlyIleAlaLeuTh 260
Qy 1110 ATTACTTAACCTGTGACAGCTCCTCGCGGGCTCCCGAGAGATTAATTAATTCGTCGCG 1169
Db 260 rLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAl 280
Qy 1170 TGGCGGGCAACTGTTTTATCCCGCCCGTGTCTACGCCAATGCGGAGCAACCGTGAA 1229
Db 280 aGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrVally 300
Qy 1230 GCTTATACATCAGTGGAGATGCTCAGCAGATAAGGTGTGCTATCCCGCCAGATAT 1289
Db 300 sLeuTyrThrSerValGluAsnAlaGlnAspLysGlyValAlaIleProHisAspIl 320
Qy 1290 CGATCTTGGTGAATTCGCGTGTGTCATTACAGATTATGACAAACACCATGAGCGAGATCG 1349
Db 320 eAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspAr 340
Qy 1350 GCCACCCCGTCGCTGGCGCATCTCGGCGCTTTTCTGTTCTCGAGCAATGATGTAAT 1409
Db 340 gProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValle 360
Qy 1410 TTGCTGTCCCTCACTGCGAGCGGAGTANGACAGTCCACTTACGGGTGTGCACTGCGCC 1469
Db 360 uTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPr 380
Qy 1470 GGTATTATCTCGGACAGCGTGACCTTGGTGAATGTTGCGATGCGCGCGCGCGGTAGC 1529
Db 380 oValTyrIleSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAl 400
Qy 1530 CCGATCGCTTGAATCGGTCCAAAGTCAACCTCGACGCGGCGCCCTCCCGACTGTTGAGA 1589
Db 400 aArgSerLeuAspTyrSerLysValThrLeuAspGlyArgProLeuProThrValGluGl 420
Qy 1590 ATATTCCAGACATCTTTGCTCCCTTCTGCTGCGAGCTCTCTTTTGGAGGCGCG 1649
Db 420 nTyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGl 440
Qy 1650 CACAAACAAAGCAGGTATCTCTATAATATATAATCTACTGTAGTACGACGATCTGAT 1709

Db 440 yThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeuIl 460
Qy 1710 TGAATAATGCTCGCGCCCATCGGTCCGCAATTCAACTATACACACGAGCTTGGGCGCG 1769
Db 460 eGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrArgLeuGlyAlaGl 480
Qy 1770 TCCGGTCCCATTTCTCGCGCGCGGTGGTTCACGCTCCCGCTCCCGCTCTGCTGGA 1829
Db 480 yProValAlaIleSerAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuGl 500
Qy 1830 GGATACCTTTGATATATCCGGGCGGCGGCACACATTTGATGACTTCTGCTGATGCGG 1889
Db 500 uAspThrPheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysAr 520
Qy 1890 CGCTTTAGGCTCCAGGCTTGTGCTTCCAGTCAACTGTGCTGAGCTCCAGCGCCTTAA 1949
Db 520 gAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLy 540
Qy 1950 AGTTAAGTGGGTAAAACTCGGAGTTG 1977
Db 540 sValLysValGlyLysThrArgGluLeu 549
RESULT 7
AAR96102
ID AAR96102 standard; protein; 540 AA.
XX
AC AAR96102;
XX
DT 16-OCT-2003 (revised)
DT 06-AUG-1996 (first entry)
XX
DE Hepatitis E virus (Mexico strain) recombinant 62K antigen.
XX
KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
KW capsid.
XX
OS Hepatitis E virus; Mexico strain.
PN WO9612807-A2.
XX
PD 02-MAY-1996.
XX
PF 23-OCT-1995; 95WO-US013703.
XX
PR 24-OCT-1994; 94US-00327952.
PR 13-OCT-1995; 95US-00542634.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Fuerst TR, Mcatee CP, Yarbrough PO, Zhang Y;
XX WPI; 1996-230608/23.
DR
PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
PT diagnostic reagents for determining HEV infection and in vaccines.
XX
PS Claim 6; Page 93-95; 125pp; English.
XX
CC DNA (AAT27110) coding for the 62K antigen (AAR96092) of hepatitis E virus
CC (HEV) Mexico strain capsid protein was cloned into baculovirus expression
CC vector pBluBacIII and recombinant 62K was expressed in Sf9 insect cells.
CC High levels of expression were obtd. and the recombinant 62K was obtd. in
CC over 98% purity. However, C-terminal processing resulted in the deletion
CC of 9 or 23 amino acids from 62K, giving 2 related polypeptide species
CC (AAR96102 and AAR96104). Similar results were obtd. with HEV Burma 62K
CC antigen (see also AAR96101 and AAR96103). Recombinant 62K represents an
CC improved antigen, in comparison to bacterial expressed proteins, for use
CC in HEV diagnostic assays, and also has excellent immunogenic properties.
XX (Updated on 16-OCT-2003 to standardise OS field)
SQ Sequence 540 AA;

Alignment Scores:

Pred. No.: 0 Length: 540
Score: 392.00 Matches: 538
Percent Similarity: 99.45% Conservative: 0
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 59.48% Indels: 3
DB: 2 Gaps: 1

US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR96102 (1-540)

```
QY 334 GCTGTGGCGCTGCCCATGACACTCACCGTCCCGGAGCTTGATCTCGCGTGCAATT 393
Db 1 AlaValAlaProAlaHisAspThrSerProValProAspValAspSerArgGlyAlaIle 20
QY 394 CTACGGCGCCAGTAAATTGCTACTTACCCCTCCACCTCCCTCTCTGCGCTCGCACT 453
Db 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaSerGlyThr 40
QY 454 AATTAGTCTGTGATGACGCCCTTAATCCGCTCTCGCGCTGCAGGACGGTACTAAT 513
Db 41 AsnLeuValLeuTyrAlaAlaProLeuAsnProProLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACTCACATTATGCCACAGAGGCTCCAATTATGACAGTACCGGTTGCCGCGTACT 573
Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThr 80
QY 574 ATCCGTTACCGGCCCTAGTGCCTAATGCAGTTGGAGGCTATGCTATATCCATTTCTTC 633
Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100
QY 634 TGGCTCAAAACACCAACCCCTACATCTGTGTGACATGAATTCCATTCTCCACTGAT 693
Db 101 TrpProGlnThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp 120
QY 694 GTCAGATCTTGTTCACCTGGCATAGCATCTGTAATGGTATCCCAAGCAGGCGCTT 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTACCGCAATCAAGGTTGGCGCTCGGTTGAGACATCTGGTGTGCTGAGGAGGAGCC 813
Db 141 HisTyrArgAsnGlnGlyTrpArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCCGGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873
Db 161 ThrSerGlyLeuValMetLeuGlyHisGlySerProValAsnSerTyrThrAsnThr 180
QY 874 CCTTATACGGTCCCTTGGCTTACTGCACTTGTGCTTAGACCTTGAGTTCGCATCTC 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200
QY 934 ACCACTGTAAACCAATACACGTGTGCTCCCTTACTCCAGCACTGCTCGTCA---CTC 989
Db 201 ThrThrCysAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
QY 990 GCCCGAGGGCGGACGGAGCTGCGGAGCTGACCAACTGCGAGCCACGAGTTCATGAA 1049
Db 221 Arg-ArgGlyAlaAspGlyThrAlaGluLeuThrThrThrAlaAlaThrArgPheMetIy 240
QY 1050 AGATCTCCACTTACCGGCTTAATGGGTAGTGAAGTCCGCGCGGAGTACGCTAAC 1109
Db 240 sAspLeuHisPheThrGlyLeuAsnGlyValGlyGluValGlyArgGlyIleAlaLeuTh 260
QY 1110 ATTACTTAACCTTGTGACAGCTCCTCGCGGGCTCCCGACAGAAATTAATTTCCTCGCG 1169
Db 260 rLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAl 280
QY 1170 TGGCGGGCAACTGTTTATTCGCGCGCGGTTGCTCAGCCAAATGGGAGCCAAACGTA 1229
Db 280 adGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrVally 300
QY 1230 GCTCTATACATCAGTGGAGATGCTCAGCAGGATAAGGTGTGTATCCCGCCACGATAT 1289
Db 300 sLeuTyrThrSerValGluAsnAlaGlnGlnAspLysGlyValAlaIleProHisAspI 320
```

```
QY 1290 CGATCTTGGTGTGATTCGGCTGTGCTCATTGAGGATTATGACAAACACCATGAGGATCG 1349
Db 320 eAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspAr 340
QY 1350 GCCCACCCCGTCCGCTGCGCCATCTCGGCTTTTCTGTTCTCCGAGCAAAATGATGTACT 1409
Db 340 gProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValle 360
QY 1410 TTGGCTGTCCCTCACCTGTCAGCGAGTATGACAGTCCACTTACGGGTCTGTCAACTGCCCC 1469
Db 360 uTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPr 380
QY 1470 GGTATTATCTCGGACAGCGTGAATGTTGGTGAATGTCGACTGGCGCGCAGCCCTAGC 1529
Db 380 ovalTyrIleSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAl 400
QY 1530 CCGATCGCTTGATCGTCCAAAGTCACCTCCAGCGCGCGCCCTCCCGACTGTTTCAGCA 1589
Db 400 aArgSerLeuAspTrpSerLysValThrLeuAspGlyA-gProLeuProThrValGluG 420
QY 1590 ATATTCCAAAGACATTTTGTGCTCCCTTCTGTCGCAAGCTCTCTTTTGGAGGCCGG 1649
Db 420 nTyrSerTyrThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaG 440
QY 1650 CACAACAAAGCAGGTTATCTTATAATTAATACTACTGCTAGTGACACGATTCCTGAT 1709
Db 440 yThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeu 460
QY 1710 TGAATATGTCGCGGCTCGGCTCGCCATTTCACTATACCATATACACGAGCTTTGGGCGCG 1769
Db 460 eGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrArgLeuGlyAlaG 480
QY 1770 TCGGTCGCGCATTTCTGGCGCGCGGTTTGGCTCCACCTCGCGCTCGCTGCTGCGA 1829
Db 480 yProValAlaIleSerAlaAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuLeuG 500
QY 1830 GGATACTTTTGTATATTCGGGCGCGCGCACACATTTGATGCTTCTGCTGCTGCTGCTGCT 1889
Db 500 uAspThrPheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysAr 520
QY 1890 CCGTTAGCGCTCCAGGTTGTCTTCCAGTCAACTGTGCTGAGCTCCAGCGCTTAA 1949
Db 520 gAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLy 540
QY 1950 A 1950
Db 540 S 540
RESULT 8
AAR96104
ID AAR96104 standard; protein; 525 AA.
XX
AC AAR96104;
XX
DT 16-OCT-2003 (revised)
DT 06-AUG-1996 (first entry)
XX
DE Hepatitis E virus (Mexico strain) recombinant 62K antigen.
XX
KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
XX capsid.
XX
OS Hepatitis E virus; Mexico strain.
XX
PN WO9612807-A2.
XX
PD 02-MAY-1996.
XX
PF 23-OCT-1995; 95WO-US013703.
XX
PR 24-OCT-1994; 94US-00327952.
```

PR 13-OCT-1995; 95US-00542634.
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX PI Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;
XX DR WPI; 1996-230608/23.
XX PS Claim 8; Page 97-98; 125pp; English.
XX CC DNA (AAT27110) coding for the 62K antigen (AAR96092) of hepatitis E virus
CC (HEV) Mexico strain capsid protein was cloned into baculovirus expression
CC vector pBluebacII and recombinant 62K was expressed in Sf9 insect cells.
CC High levels of expression were obtd. and the recombinant 62K was obtd. in
CC over 95% purity. However, C-terminal processing resulted in the deletion
CC of 9 or 23 amino acids from 62K, giving 2 related polypeptide species
CC (AAR96102 and AAR96104). Similar results were obtd. with HEV Burma 62K
CC antigen (see also AAR96101 and AAR96103). Recombinant 62K represents an
CC improved antigen, in comparison to bacterial expressed proteins, for use
CC in HEV diagnostic assays, and also has excellent immunogenic properties.
CC (Updated on 16-OCT-2003 to standardise OS field)
XX SQ Sequence 525 AA;

Alignment Scores:
Pred. No.: 0 Length: 525
Score: 377.00 Matches: 523
Percent Similarity: 99.43% Conservative: 0
Best Local Similarity: 99.43% Mismatches: 1
Query Match: 57.21% Indels: 3
DB: 2 Gaps: 1

US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR96104 (1-525)

QY 334 GCTGTGGCGCTGCCATGACCTACCTCCGTCGGAGCTTGATCTCCGGTGCAT 393
Db 1 AlaValAlaProAlaHisAspThrSerProValProAspValAspSerArgGlyAlaIle 20

QY 394 CTAGCGCGCCAGTAAATTTGTCTACTTCAACCCCTGCATCCTCTGTGGCTCTGGCACT 453
Db 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaSerGlyThr 40

QY 454 AATTAGTCTGTATGACGCCCCCTTAATCCGGCTCTGGCGCTGCGAGGAGCTACTAAT 513
Db 41 AsnLeuValLeuTyrAlaAlaProLeuAsnProProLeuProLeuGlnAspGlyThrAsn 60

QY 514 ACTCACATTATGCCACAGAGGCTCCAATTATGACAGTACCGGTTGCCGCGTACT 573
Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaAlaArgAlaThr 80

QY 574 ATCCGTTACCGGCGCCCTAGTGCCTAAATGACAGTTCGAGGCTATGCTATATCCATTCTTTC 633
Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100

QY 634 TGGCCTCAACACACCAACCCCTACATCTGTTGACATGATTAATCCATTACTTCCACTGAT 693
Db 101 TrpProGlnThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp 120

QY 694 GTCAGGATTCTGTTCACCTGGCATAGCATCTGAATTGGTTCATCCCAAGCGAGCGCTT 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGlnLeuValIleProSerGluArgLeu 140

QY 754 CACTACCGCAATCAAGGTTGGCGCTCGGTTGAGACATCTGTGTTGCTGAGGAGGAAGCC 813
Db 141 HisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGluGluGluAla 160

QY 814 ACCTCCGGTCTGTGTCATGTTATGCATACATGGCTCTCCAGTTAACTCCTATACCAATACC 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThr 180

RESULT 9

QY 874 CCTTATACCGGTGCCCTTGGCTTACTGCAGCTTGGCTTAGAGCTTGAGTTTCGCAATCTC 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200

QY 934 ACCACCTGTAAACACCAATACACCGTGTGTCCTCCGTTACTCCAGCACTGCTGCTCA---CTC 989
Db 201 ThrThrCysAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220

QY 990 GCGCCGAGGGCGGACCGGAGCTCGGAGCTGACCACTGCAGCCACCGAGTTTCATGAA 1049
Db 221 Arg-ArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLeu 240

QY 1050 AGATCTCCACTTACCGGCTTAATGGGTAGGTGAAGTGGCGCGCGGATAGCTCTTAAC 1109
Db 240 sAspLeuHisPheThrGlyLeuAsnGlyValGlyGluValGlyArgGlyIleAlaLeuTh 260

QY 1110 ATTACTTTAACTTGTCTGACACCGTCTCTCGCGGGCTCCCGACAGAAATAATTTTCGTGGC 1169
Db 260 rLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAl 280

QY 1170 TGGCGGGCAACTGTTTATTCCCGCCCGGTTGCTCAGCCAAATGGCGAGCCAAACCGTGA 1229
Db 280 aGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLy 300

QY 1230 GCTCTATACATCAGTCGAGAAATGCTCAGCAGGATAAGGGTGTGCTATCCCCACGATAT 1289
Db 300 sLeuTyrThrSerValGluAsnAlaGlnAspGlyValAlaIleProHisAspIle 320

QY 1290 CGATCTTGGTGATTGCGTGTGTCATTTCAGGATTATGACAACCGCATGAGCAGGATCG 1349
Db 320 eAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspAr 340

QY 1350 GCCACCCCGTGGCTCGGCCATCTCGGCTTTTCTGTTCTCGAGCAAAATGATGACT 1409
Db 340 gProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLe 360

QY 1410 TTGCTCTCCCTCACTCAGCGAGTATGACCTCCACTTACGGGTGCTCAACTGGCC 1469
Db 360 uTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPr 380

QY 1470 GGTATTATCTCGACAGCGTGACTTTGGTGAATGTTGCGACTGGCGCGAGGCGGTAGC 1529
Db 380 oValTyrIleSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAl 400

QY 1530 CCGATCGCTGACTGTCGCAAGTCAACCTCGAGGGCGGCGCCCTCCGACTGTTGAGCA 1589
Db 400 aArgSerLeuAspTrpSerLysValThrLeuAspGlyArgProLeuProThrValGluGl 420

QY 1590 ATATTCCACAGACATCTTGTGCTCCCGCTTCGTGGCAAGCTCTCTTTTGGGAGGCGCG 1649
Db 420 nTyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheThrGluAlaGl 440

QY 1650 CACAACAAAAGAGGTTATCTTTATATAATACTACTGCTAGTGACCAAGATTCTGAT 1709
Db 440 yThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeuIl 460

QY 1710 TGAATATGCTCGCGCCATCGGTCGCCATTTCACTATACCAAGGCTTGGGGCGCG 1769
Db 460 eGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrArgLeuGlyAlaGl 480

QY 1770 TCCGCTCGCATTTCTGGCGCGCGGTTTGGCTCCACGCTCCGCGCTGCTGCTGGA 1829
Db 480 yProValAlaIleSerAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuLeuGl 500

QY 1830 GGATCTCTTTGATTATCCGGGGCGGCGCACATTTGATGACTTCTGCGCTGAATCCG 1889
Db 500 uAspThrPheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysAr 520

QY 1890 CGCTTAGGCTCCAG 1905
Db 520 gAlaLeuGlyLeuGln 525

AAR96094
ID AAR96094 standard; protein; 327 AA.
AC AAR96094;
XX
DT 16-OCT-2003 (revised)
DT 06-AUG-1996 (first entry)
XX
DE Hepatitis E virus (Mexico strain) SG3 antigen.
XX
KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
KW diagnosis; antigen.
XX
OS Hepatitis E virus; Mexico strain.
XX
PN WO9612807-A2.
XX
PD 02-MAY-1996.
XX
PF 23-OCT-1995; 95WO-US013703.
XX
PR 24-OCT-1994; 94US-00327952.
PR 13-OCT-1995; 95US-00542634.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;
XX
DR WPI; 1996-230608/23.
DR N-PSDB; AAT27112.
XX
PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
PT diagnostic reagents for determining HEV infection and in vaccines.
XX
PS Disclosure; Page 87-88; 125pp; English.
XX
CC SG3 antigen (AAR96094) comprises the C-terminal 327 amino acids of
CC hepatitis E virus (HEV) Mexico strain capsid protein (see also AAR96090).
CC It is obtd. inserting encoding DNA (AAT27112) into pBluescript and
CC ligating the cloned insert into vector pGEX for prodn. of SG3 in E. coli
CC cells. SG3 shows low antigenicity compared with capsid 62K antigen (see
CC also AAR96091-92 and AAR96101-04) from HEV Burma and Mexico strains that
CC are produced using a baculovirus expression system. (Updated on 16-OCT-
CC 2003 to standardise OS field)
XX
SQ Sequence 327 AA;
XX
Alignment Scores:
Pred. No.: 9,85e-308 Length: 327
Score: 327.00 Matches: 327
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.62% Indels: 0
DB: 2 Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR96094 (1-327)
QY 997 GGGGCGGAGGACTGCGGAGCTGACCAACTGCAGCCACCAAGTTCATGAAGATCTC 1056
Db 1 GlyAlaAspGlyThrAlaGluLeuThrThrThrAlaAlaThrArgPheMetLysAspLeu 20
QY 1057 CACTTTACCGGCTTAATGGGTAGGTGAAGTCGGCGGGGATAGCTCTAACATTACTT 1116
Db 21 HisPheThrGlyLeuAsnGlyValGlyGluValGlyArgGlyLeuAlaLeuThrLeuLeu 40
QY 1117 AACCTTGCTGACACGCTCTCGGCGGGCTCCGACAGAAATTAATTCGTGCGCTGGCGG 1176
Db 41 AsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuLeuLeuSerAlaGlyGly 60
QY 1177 CAACCTGTTTTATTCCCGCGGCTTGTCTCAGCAATGGGAGCCCAACCGTGAAGCTCTAT 1236
Db 61 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 80

QY 1237 ACATCAGTGGGAATGCTCAGCAGGATAAGGGTGTGTGCTATCCCCACGATATCGATCTT 1296
Db 81 ThrSerValGluAsnAlaGlnAspLysGlyValAlaAlaProHisAspLeuLeu 100
QY 1297 GGTGATTCGGTGTGTCATTGAGATTAAGCAACCCAGCATGACAGGATCGGCCACCC 1356
Db 101 GlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 120
QY 1357 CGGTGCGCTGGCCATCTCGGCTTTTCTCTCCGAGCAAAATGATGTACTTTGGCTG 1416
Db 121 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuThrLeu 140
QY 1417 TCCCTCACTGCAGCGGAGTATGACCACTCCACTTACGGGTCTCGTCAACTGGCCCGTTTAT 1476
Db 141 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 160
QY 1477 ATCTCGACAGGTGACTTTGGTGAATGTTGGGACTGGCGGCGCAGCGGTAGCCGATCG 1536
Db 161 IleSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValaAlaArgSer 180
QY 1537 CTTGACTGCTCCAAAGTCACCTCGACGGGGCGCCCTCCCGACTGTTGAGCAATATTC 1596
Db 181 LeuAspTyrSerLysValThrLeuAspGlyArgProLeuProThrValGluGlnTyrSer 200
QY 1597 AAGACATTTCTTTGTGCTCCCTTCGTGGCAAGCTCTCTTTTGGAGCGCCGACACA 1656
Db 201 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 220
QY 1657 AAAGCAGGTATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1716
Db 221 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsn 240
QY 1717 GCTGCGCGCATCGGCTCGCCATTTCAACCTATACACAGGCTTGGGGCGGTCGGGTC 1776
Db 241 AlaAlaGlyHisArgValAlaIleSerThrThrThrArgGluGlyAlaGlyProVal 260
QY 1777 GCCATTTCTGGCGCGCGGCTTTGGCTCCACGCTCCGCTCGCTGCTGAGGATACT 1836
Db 261 AlaIleSerAlaAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuLeuGluAspThr 280
QY 1837 TTTGATTAATCGGGCGGGCGGCGACACATTTGATGACTTCTGCTGCTGATCGCGCTTTA 1896
Db 281 PheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysArgAlaLeu 300
QY 1897 GGCCTCCAGGTTGTGCTTCCAGTCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAG 1956
Db 301 GlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysValLys 320
QY 1957 GTGGTAAACTCGGAGTTG 1977
Db 321 ValGlyLysThrArgGluLeu 327
RESULT 10
AAW35821
ID AAW35821 standard; protein; 327 AA.
XX
AC AAW35821;
XX
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 26-FEB-1998 (first entry)
XX
DE Hepatitis E virus Mexico strain peptide antigen SG3.
XX
KW Hepatitis E virus; Burma; Mexico; immunoassay; peptide antigen; antibody;
KW diagnosis; HEV.
XX
OS Hepatitis E virus; - Mexico strain.
XX
PN US5686239-A.
XX
PD 11-NOV-1997.
XX

QY 490 CTGCGCTGACGAGCGGTACTATCACTATGAGCCACAGAGCGCTCCAAATTATGCA 549
Db |||||||
53 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 72
QY 550 CAGTACCGGGTTGCCGCGCTACTATCCGTTACCGGCCCCCTAGTGCCTTAATGCAGTTGGA 609
Db |||||||
73 GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 92
QY 610 GGCTATGCTATATCATCTCTCTTCTGGCTCAAAACACCAACCCCTACATCTCTGTTGAC 669
Db |||||||
93 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 112
QY 670 ATGAATTCATTAATCTTCCACTGATGTCAGGATTCCTGTTCAACCTGGCATAGCATCTGAA 729
Db |||||||
113 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 132
QY 730 TTGCTCATCCCAAGCGCGCTTCACTACCGCATCAAGGTTGCGCTCGGTTGAGACA 789
Db |||||||
133 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 152
QY 790 TCTGCTGTGCTGAGGAGGAGCCACCTCCGCTCTTGTGCTGATGTTATGCATATGCTCT 849
Db |||||||
153 SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 172
QY 850 CCAGTTAACTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGGACTTTGCC 909
Db |||||||
173 ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla 192
QY 910 TTAGAGCTTGAGTTTCGCAATCTCACC 936
Db |||||||
193 LeuGluLeuGluPheArgAsnLeuThr 201

RESULT 13

AAR70323
ID AAR70323 standard; protein; 660 AA.

XX AC AAR70323;

XX XX 25-MAR-2003 (revised)

DT 28-SEP-1995 (first entry)

XX XX Hepatitis E virus ORF2 protein.

XX HEV; ORF2; antigen; vaccine; immunogen.

XX OS Hepatitis E virus.

XX XX WO9508632-A1.

XX PD 30-MAR-1995.

XX PF 23-SEP-1994; 94WO-AU000572.

XX PR 24-SEP-1993; 93AU-00001423.

XX PR 15-DEC-1993; 93AU-00002964.

XX XX (MACF-) MACFARLANE BURNET CENT MEDICAL.

XX PI Anderson DA, Locarnini SA, Torresi J, Li F, Hui Z;

XX DR WPI; 1995-139601/18.

XX DR N-PSDE; AAQ86592.

XX PT Antigens of hepatitis E virus (HEV) - selectively immuno-reactive to

XX convalescent and/or acute phase circulating antibodies to HEV.

XX PS Disclosure; Page 39-41; 78pp; English.

XX CC RNA from an HEV strain isolated in the Xinjiang region of China was

XX CC subjected to RT-PCR to obtain fragments corresp. to ORF2 (given in

XX CC AAQ86592) and ORF3 (AAQ86593) that encoded antigenic proteins (AAR70323-

XX CC 24, respectively). DNA fragments were manipulated into pGEX vectors for

XX CC production of GST fusion proteins in E. coli. Applications include HEV

CC infection diagnosis, therapy and vaccine development. (Updated on 25-MAR-

CC 2003 to correct PN field.)

SQ Sequence 660 AA;

Alignment Scores:

Pred. No.: 4,23e-135 Length: 660
Score: 149.00 Matches: 149
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.61% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR70323 (1-660)

QY 490 CTGCGCTGACGAGCGGTACTATCACTATGAGCCACAGAGCGCTCCAAATTATGCA 549

Db |||||||

164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183

QY 550 CAGTACCGGGTTGCCGCGCTACTATCCGTTACCGGCCCCCTAGTGCCTTAATGCAGTTGGA 609

Db |||||||

184 GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203

QY 610 GGCTATGCTATATCATCTCTCTTCTGGCTCAAAACACCAACCCCTACATCTCTGTTGAC 669

Db |||||||

204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223

QY 670 ATGAATTCATTAATCTTCCACTGATGTCAGGATTCCTGTTCAACCTGGCATAGCATCTGAA 729

Db |||||||

224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243

QY 730 TTGCTCATCCCAAGCGCGCTTCACTACCGCATCAAGGTTGCGCTCGGTTGAGACA 789

Db |||||||

244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 263

QY 790 TCTGCTGTGCTGAGGAGGAGCCACCTCCGCTCTTGTGCTGATGTTATGCATATGCTCTCT 849

Db |||||||

264 SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283

QY 850 CCAGTTAACTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGGACTTTGCC 909

Db |||||||

284 ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla 303

QY 910 TTAGAGCTTGAGTTTCGCAATCTCACC 936

Db |||||||

304 LeuGluLeuGluPheArgAsnLeuThr 312

RESULT 14

AAR91814

ID AAR91814 standard; protein; 660 AA.

XX AC AAR91814;

XX XX 26-NOV-1996 (first entry)

DT 26-NOV-1996 (first entry)

XX XX Hepatitis E virus strain SAR-55 ORF-2.

XX KW Hepatitis E virus; HEV; SAR-55 strain; enteric transmission;

XX KW structural region; antigen; detection; antibody; vaccine; immunisation;

XX KW infection.

XX OS Hepatitis E virus.

XX PN WO9610590-A2.

XX PD 11-APR-1996.

XX PF 03-OCT-1995; 95WO-US013102.

XX PR 03-OCT-1994; 94US-00316765.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX XX

PI Tsarev SA, Emerson SU, Purcell RH;
XX WPI, 1996-209320/21.
DR N-PSDB; AAT27394.
XX
PT Isolated and purified hepatitis E virus strain SAR-55 DNA - encodes
PT antigenic protein useful in diagnosis, prophylaxis and treatment of
PT hepatitis E virus infection.
XX
XX Claim 6; Page 13-15; 121pp; English.
PS
XX The present sequence is the protein prod. of ORF-2 from the hepatitis E
CC virus (HEV) strain SAR-55, which was implicated in an enterically
CC transmitted non-A, non-B hepatitis in Pakistan. The protein encoded by
CC the structural region of the virus (i.e. ORF-2), which is capable of
CC forming HEV like particles, is useful for the detection of HEV antibodies
CC (pref. IgG or IgM) in blood, plasma, sera, cerebrospinal fluid, tissue,
CC urine or pleural fluid. The protein, and anti-HEV antibodies generated
CC using the protein, can also be used in vaccines for immunising an animal
CC against HEV infection. The protein is identified as a band of greater
CC than 50 kD following SDS-PAGE of cell lysates of insect cells infected
CC with a HEV ORF-2 cong. baculovirus, i.e. the claimed recombinant
CC expression vectors pPIC9-1779, -1780 and -1781
XX
SQ Sequence 660 AA;

Alignment Scores:
Pred. No.: 4,238-135 Length: 660
Score: 149.00 Matches: 149
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.61% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR91814 (1-660)

QY 490 CTGCGCGCTGCAGAGCGTACTAATCTACATTATGGCCACAGAGCGCTCCCAATTATGCA 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183

QY 550 CAGTACCGGGTGGCCCGCGCTACTATCCGTTACCGGCCCTAGTGGCTTAATCGAGTTGA 609
Db 184 GlnTyrArgValAlaAargAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203

QY 610 GCCTATGCTATATTCATTCTTTCTGGCCTCAAAACACACACCCCTACATCTGTTGAC 669
Db 204 GlyTyrAlaIleSerIleSerPheThrProGlnThrThrThrProThrSerValAsp 223

QY 670 ATGAATTCCATTACTTCCACTGATGTCAGGATTCTTGTTCACCTGGCATAGCATCTGAA 729
Db 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243

QY 730 TTGGTCAATCCCAAGGAGCGCTTACTACCGCAATCAAGTTGGCGCTCGGTTGAGACA 789
Db 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrPargSerValGluThr 263

QY 790 TCTGGTGTGTCGAGGAGGAGCGACCTCCGGTCTTGTTCATGTTATGCATATGCTGCTCT 849
Db 264 SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283

QY 850 CCAGTTAACTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGAGACTTTGCC 909
Db 284 ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla 303

QY 910 TTAGACTTGAGTTTCGGCAATCTCACC 936
Db 304 LeuGluLeuGluPheArgAsnLeuThr 312

RESULT 15
AAW76369
ID AAW76369 standard; protein; 660 AA.
XX
AC AAW76369;

XX 03-DEC-1998 (first entry)
DT
XX Hepatitis E virus hollow particle protein #2.
DE
XX Hollow particle protein; virus; antibody; detection; immunoassay;
KW infection.
KW
OS Hepatitis virus.
XX
PN JP10234383-A.
XX
PD 08-SEP-1998.
XX
PF 28-FEB-1997; 97JP-00062445.
XX
PR 28-FEB-1997; 97JP-00062445.
XX
PA (DENK-) DENKA SEIKEN KK.
XX (KOKU-) KOKURITSU YODO BISEI KENKYUSHO.
XX
DR WPI: 1998-535037/46.
DR N-PSDB; AAV61888.
XX
PT Hepatitis E virus hollow particle poly:peptide(s) and nucleic acids
PT encoding it - useful for more accurate detection of HEV in samples, using
PT immuno-assays and nucleic acid hybridisation.
XX
XX Claim 13; Page 24-26; 29pp; Japanese.
PS
CC This sequence represents a Hepatitis E viral hollow particle protein.
CC This polypeptide can be used to raise antibodies to detect HEV infection
CC in samples, e.g. by immuno-assay based techniques, and the nucleic acid
CC can be used for the same in nucleic acid hybridisation assays. The
CC polypeptides and nucleic acids allow more accurate detection of HEV than
CC previously possible
XX
SQ Sequence 660 AA;

Alignment Scores:
Pred. No.: 4,238-135 Length: 660
Score: 149.00 Matches: 149
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.61% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAW76369 (1-660)

QY 490 CTGCGCGCTGCAGAGCGTACTAATCTACATTATGGCCACAGAGCGCTCCCAATTATGCA 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183

QY 550 CAGTACCGGGTGGCCCGCGCTACTATCCGTTACCGGCCCTAGTGGCTTAATCGAGTTGA 609
Db 184 GlnTyrArgValAlaAargAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203

QY 610 GCCTATGCTATATTCATTCTTTCTGGCCTCAAAACACACACCCCTACATCTGTTGAC 669
Db 204 GlyTyrAlaIleSerIleSerPheThrProGlnThrThrThrProThrSerValAsp 223

QY 670 ATGAATTCCATTACTTCCACTGATGTCAGGATTCTTGTTCACCTGGCATAGCATCTGAA 729
Db 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243

QY 730 TTGGTCAATCCCAAGGAGCGCTTACTACCGCAATCAAGTTGGCGCTCGGTTGAGACA 789
Db 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrPargSerValGluThr 263

QY 790 TCTGGTGTGTCGAGGAGGAGCGACCTCCGGTCTTGTTCATGTTATGCATATGCTGCTCT 849
Db 264 SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283

QY 850 CCAGTTAACTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGAGACTTTGCC 909
Db 284 ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla 303

QY 910 TTAGACTTGAGTTTCGGCAATCTCACC 936
Db 304 LeuGluLeuGluPheArgAsnLeuThr 312

RESULT 15
AAW76369
ID AAW76369 standard; protein; 660 AA.
XX
AC AAW76369;

QY	850	CCAGTAACTCGTATACCAATAACCCCTTATACCGTGCCTTGCCTTACTCGACTTTGCC	909
Db <td>284 <td>ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuAspPheAla <td>303</td> </td></td>	284 <td>ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuAspPheAla <td>303</td> </td>	ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuAspPheAla <td>303</td>	303
QY <td>910 <td>TTAGAGCTTGAGTTTCGCAATCTCACC</td> <td>936</td> </td>	910 <td>TTAGAGCTTGAGTTTCGCAATCTCACC</td> <td>936</td>	TTAGAGCTTGAGTTTCGCAATCTCACC	936
Db <td>304 <td>LeuGluLeuGluPheArgAsnLeuThr</td> <td>312</td> </td>	304 <td>LeuGluLeuGluPheArgAsnLeuThr</td> <td>312</td>	LeuGluLeuGluPheArgAsnLeuThr	312

Search completed: August 16, 2004, 13:48:11
Job time : 114.682 secs

Blank sheet

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.21% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VHHWH2 (1-660)

```
QY 490 CTGCGCGTGCAGGACGGTACTAATCACTTACATTAATGGCCACAGAGCGCTCCAAATTATGCA 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183
QY 550 CAGTACCGGGTTGCCGGCTACTATCCGTTCACGGCCCTAGTGCCTTAATGCAGTTGGA 609
Db 184 GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203
QY 610 GGCTATGCTATATCCATTCTTCTTGGCTCAAAACACCAACCCCTACATCTGTTGAC 669
Db 204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCCTACTTCTCCACTGTGTGAGGATTCTTGTTCACCTGGCATGCACTCTGAA 729
Db 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
QY 730 TTGCTATCCCAAGCAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCGGTTGACACA 789
Db 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThr 263
QY 790 TCTGGTCTCTGAGGAGGAGCCACCTCCGGTCTTGTCAATGTTATGCATACATGGCTCT 849
Db 264 SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283
```

RESULT 3

C44212
structural protein 1 - hepatitis E virus (strain Mexico)
C;Species: hepatitis E virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C;Accession: C44212
R;Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.
Virology 191, 550-558, 1992
A;Title: Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HE)
A;Reference number: A44212; MUID:91079857; PMID:11448913
A;Molecule type: Genomic RNA
A;Residues: 1-123 <HUA>
A;Cross-references: GB:M74506; NID:g330017; PIDN:AAA45731.1; PID:g330019
C;Superfamily: hepatitis E virus structural protein 1
C;Keywords: structural protein

Alignment Scores:
Pred. No.: 5.5e-101 Length: 123
Score: 109.00 Matches: 109
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.54% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x C44212 (1-123)

```
QY 2 TGGCCCTAGCCCTCTTTGCTGTGTTCCTCTGTTTCTGCCTATGTTGCCCGCGCCAC 61
Db 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGACCGGTACGCGCTCTGGCCGCGCTGTGGCGGCGCAGCGGGTACCGCGGTGGTT 121
Db 35 ArgProValSerArgLeuAlaValAlaValGlyGlyAlaAlaValProAlaValVal 54
QY 122 TCTGGGTACCGGGTGTATCTCAGCCCTTCGCAATCCCCCTATATCATCAACCAACC 181
Db 55 SerGlyValThrGlyLeuLeuLeuSerProSerGlnSerProIlePheIleGlnProThr 74
QY 182 CTTTGGCCCGCAGCGTTCGCGCTCGCGGTCTGACCTCGCTTCGCCAACACGACCC 241
Db 75 ProLeuProGlnThrLeuProLeuArgProGlyLeuAspLeuAlaPheAlaAsnGlnPro 94
```

```
QY 242 GGCCACTTGGCTCCACTTGGCGAGATCAGGCCACAGCCCTCGCTCCCTCCGTCGCC 301
Db 95 GlyHisLeuAlaProLeuGlyGluIleArgProSerAlaProLeuProProValAla 114
QY 302 GACCTGCCACACAGCGGGGCTGCGGCGC 328
Db 115 AspLeuProGlnProGlyLeuArgArg 123
```

RESULT 4

VHHWHE
structural protein 1 - hepatitis E virus (strain Burma)
C;Species: hepatitis E virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 23-Jul-1999
C;Accession: B40778; A40236
R;Tam, A.W.; Smith, M.M.; Guerra, M.E.; Huang, C.C.; Bradley, D.W.; Fry, K.E.; Reyes, G.F.
Virology 185, 120-131, 1991
A;Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length vir
A;Reference number: A40778; MUID:92024067; PMID:1926770
A;Accession: B40778
A;Molecule type: genomic RNA
A;Residues: 1-123 <TAM>
A;Cross-references: GB:M73218; NID:g330023; PIDN:AAA45735.1; PID:g330025
R;Ray, R.; Jameel, S.; Manivel, V.; Ray, R.
Virology 189, 359-362, 1992
A;Title: Indian hepatitis E virus shows a major deletion in the small open reading frame.
A;Reference number: A40236; MUID:92295577; PMID:1534953
A;Accession: A40236
A;Molecule type: genomic RNA
A;Residues: 1-57,'p',59-102,'p',104-123 <RAY>
C;Superfamily: hepatitis E virus structural protein 1
C;Keywords: structural protein

Alignment Scores:
Pred. No.: 1.9e-52 Length: 123
Score: 61.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.26% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VHHWHE (1-123)

```
QY 2 TGGCCCTAGCCCTCTTTGCTGTGTTCCTCTGTTTCTGCCTATGTTGCCCGCGCCAC 61
Db 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGACCGGTACGCGCTCTGGCCGCGCTGTGGCGGCGCAGCGGGTACCGCGGTGGTT 121
Db 35 ArgProValSerArgLeuAlaValAlaValGlyGlyAlaAlaValProAlaValVal 54
QY 122 TCTGGGTACCGGGTGTATCTCAGCCCTTCGCAATCCCCCTATATCATCAACCAACC 181
Db 55 SerGlyValThrGlyLeuLeuLeuSerProSerGlnSerProIlePheIleGlnProThr 74
QY 182 CCT 184
Db 75 Pro 75
```

RESULT 5

B40236
structural protein 1 - hepatitis E virus (strain Indian)
C;Species: hepatitis E virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: B40236
R;Ray, R.; Jameel, S.; Manivel, V.; Ray, R.
Virology 189, 359-362, 1992
A;Title: Indian hepatitis E virus shows a major deletion in the small open reading frame.
A;Reference number: A40236; MUID:92295577; PMID:1534953
A;Accession: B40236
A;Molecule type: genomic RNA
A;Residues: 1-41 <RAY>
C;Superfamily: hepatitis E virus structural protein 1

C;Keywords: structural protein

Alignment Scores:
Pred. No.: 0.000778 Length: 41
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.97% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x B40236 (1-41)

QY 2 TCGGCGCTAGGCTCTTTTCGCTGTTCTCTCTGTTTC 40

Db 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPhe 27

RESULT 6

T36459

hypothetical protein SCF43A.36 - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T36459

R;Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1999

A;Reference number: Z21598

A;Accession: T36459

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-266 <SEE>

A;Cross-references: EMBL:AL096837; PIDN:CAB48923.1; GSPDB:GN00070; SCOEDB:SCF43A.36

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEDB:SCF43A.36

Alignment Scores:
Pred. No.: 6.51 Length: 266
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x T36459 (1-266)

QY 76 AGCGCTGACCGCTCGCTGCGCGGCA 50

Db 100 ThrAlaAspArgSerValAlaArgAla 108

RESULT 7

G69580

acetyl-CoA carboxylase (EC 6.4.1.2), carboxyltransferase alpha chain - Bacillus subtilis

C;Species: Bacillus subtilis

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C;Accession: G69580

R;Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

A.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Puma, S.; Galizzi, A.; Gallen

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Konningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinols,

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel

Y, M.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: G69580

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-325 <KUN>

A;Cross-references: GB:Z99118; GB:AL009126; NID:g2635200; PIDN:CAB14880.1; PID:g2635385

A;Experimental source: strain 168

C;Genetics:

A;Gene: accA

C;Function:

A;Pathway: fatty acid biosynthesis

C;Superfamily: acetyl-CoA carboxylase, carboxyltransferase alpha chain

C;Keywords: fatty acid biosynthesis, ligase

Alignment Scores:
Pred. No.: 6.31 Length: 325
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x G69580 (1-325)

QY 294 GGAGGACGGAGGGCGCTGGCGCTG 268

Db 199 GlyGlySerGlyGlyAlaLeuGlyLeu 207

RESULT 8

H58208

protamine II - black rat snake (fragment)

C;Species: Elaphe obsoleta quadrivittata

C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 07-May-1999

C;Accession: H58208

R;Hunt, J.G.; Kasinsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp, D.

J. Biol. Chem. 271, 23547-23557, 1996

A;Title: Protamines of reptiles.

A;Reference number: A58208; MUID:963994458; PMID:8798564

A;Accession: H58208

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-41 <HUN>

C;Superfamily: sperm histone

Alignment Scores:
Pred. No.: 88.5 Length: 41
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x H58208 (1-41)

QY 76 TCTGGCGCGCTGTCGGCGGCGGC 99

Db 13 SerGlyArgArgGlyArgArg 20

RESULT 9

S19708

hypothetical protein C - Herpetosiphon aurantiacus

C;Species: Herpetosiphon aurantiacus

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Jan-1995

C;Accession: S19708

R;Brdmann, D.; Duesterhoeft, A.; Kroeger, M.

Eur. J. Biochem. 202, 1247-1256, 1991

A;Title: Cloning and molecular characterization of the HgiCI restriction/ modification

A;Reference number: S19706; MUID:92111503; PMID:1662609

A;Accession: S19708

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-74 <KRO>

A;Cross-references: EMBL:X55138

Alignment Scores:
Pred. No.: 81 Length: 74
Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x S19708 (1-74)

Qy 827 ACAAGACGGAGTGGCTTCCTCC 804

Db 64 ThrArgProGluValAlaSerSer 71

RESULT 10

C61547

hypothetical protein 1 (px region) - human T-cell lymphotropic virus type 1 (isolate HAM

C;Species: human T-cell lymphotropic virus type 1, HTLV-1

C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 08-Oct-1999

C;Accession: C61547

R;Tsujimoto, A.; Teruuchi, T.; Imamura, J.; Shimotohno, K.; Miyoshi, I.; Miwa, M.

Mol. Biol. Med. 5, 29-42, 1988

A;Title: Nucleotide sequence analysis of a provirus derived from HTLV-1-associated myelo

A;Reference number: A61547; MUID:88232270; PMID:2897612

A;Accession: C61547

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-99 <TSU>

A;Cross-references: GB:M37301; NID:G541634; PIDN:AAA45390.1; PID:G541637

Alignment Scores:
Pred. No.: 77.5 Length: 99
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x C61547 (1-99)

Qy 22 CTGTGTTCCTCTGTCTTCGCCT 45

Db 52 LeuLeuPheLeuLeuPheLeuPro 59

RESULT 11

C90345

hypothetical protein SS01819 [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C;Accession: C90345

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayes, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: C90345

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-104 <KUR>

A;Cross-references: GB:AE006641; NID:gl3815078; PIDN:AAK42018.1; GSPDB:GN00155

C;Genetics:

A;Gene: SS01819

Alignment Scores:
Pred. No.: 76.9 Length: 104
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x C90345 (1-104)

Qy 1115 TTAACCTTGTCACACGCTCCTCG 1138

|||||

Db 16 LeuThrLeuLeuThrArgSerSer 23

RESULT 12

D75495

hypothetical protein - Deinococcus radiodurans (strain R1)

C;Species: Deinococcus radiodurans

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000

C;Accession: D75495

R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I

, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Mal

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: D75495

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-142 <WHI>

A;Cross-references: GB:AE001921; GB:AE000513; NID:G6458330; PIDN:AAF10219.1; PID:G6458341

A;Experimental source: strain R1

C;Genetics:

A;Gene: DR0638

A;Map position: 1

C;Superfamily: Deinococcus radiodurans hypothetical protein DR0638

Alignment Scores:
Pred. No.: 73.4 Length: 142
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x D75495 (1-142)

Qy 305 GGTCGGGACGGAGCGAGCGAG 282

|||||

Db 97 GlyArgArgGluAlaAlaGlu 104

RESULT 13

JQ1865

hypothetical 15.1K protein - bovine adenovirus 3

C;Species: Mastadenovirus bos3 (bovine adenovirus 3)

C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999

C;Accession: JQ1865

R;Mittal, S.K.; Prevec, L.; Babiuk, L.A.; Graham, F.L.

J. Gen. Virol. 73, 3295-3300, 1992

A;Title: Sequence analysis of bovine adenovirus type 3 early region 3 and fibre protein 3

A;Reference number: PQ0499; MUID:93107871; PMID:1459367

A;Accession: JQ1865

A;Molecule type: DNA

A;Residues: 1-142 <MIT>

A;Cross-references: GB:M12928

A;Experimental source: strain WBR-1

Alignment Scores:
Pred. No.: 73.4 Length: 142
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x JQ1865 (1-142)

Qy 1022 CCACAACTGCAGCCACCAGGTCA 1045

|||||

Db 57 ProGlnLeuGlnProProGlySer 64

RESULT 14

A46181

pX-rer-orf I (alternatively spliced) - human T-cell lymphotropic virus type 1

C;Species: human T-cell lymphotropic virus type 1, HTLV-1

C;Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000
C;Accession: A46181; B46181
R;Korallnik, I.J.; Gessain, A.; Klotman, M.E.; Lo Monaco, A.; Berneman, Z.N.; Franchini,
Proc. Natl. Acad. Sci. U.S.A. 89, 8813-8817, 1992
A;Title: Protein isoforms encoded by the pX region of human T-cell leukemia/lymphotropic
A;Reference number: A46181; MUID:92409607; PMID:1528897

A;Accession: A46181
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-152 <KOR>
A;Note: sequence extracted from NCBI backbone (NCBIP:114304)
A;Accession: B46181
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 54-152 <K02>
A;Note: sequence extracted from NCBI backbone (NCBIP:114306)

Alignment Scores:
Pred. No.: 72.7 Length: 152
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x A46181 (1-152)

QY 22 CTGTTGTCCTCTGTTCTGCCT 45
DB 105 LeuLeuPheLeuLeuPheLeuPro 112

RESULT 15

D87322
hypothetical protein CC0590 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: D87322
R;Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
R. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: D87322
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-153 <STO>
A;Cross-references: GB:AE005673; NID:gi3421790; PIDN:AAK22576.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC0590

Alignment Scores:
Pred. No.: 72.6 Length: 153
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x D87322 (1-153)

QY 313 GCCGGGCTGCGGCGCTGACGGCT 336
DB 109 AlaGlyAlaAlaAlaLeuThrAla 116

Search completed: August 16, 2004, 13:57:47
Job time : 39.2263 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:36:13 ; Search time 18.2362 Seconds
(without alignments)
11307.070 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096

Perfect score: 659

Sequence: 1 ATGGCCCTAGGCTCTTTT.....GTAAACTCGGAGTTGTAG 1980

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Word size: 1

Total number of hits satisfying chosen parameters: 283186

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09851410/runat_16082004_125716_9625/app_query.fasta_1.4238
-DB=SwissProt 42 -QFMT=fastan -SUFFIX=oli.rsp -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09851410@cgn 1 1.43 @runat_16082004_125716_9625 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	659	100.0	659	VST2_HEVME	Q03500 hepatitis e
2	149	22.6	660	VST2_HEVFA	P33426 hepatitis e
3	124	18.8	485	VST2_HEVRH	Q00270 hepatitis e
4	124	18.8	660	VST2_HEVMY	Q04611 hepatitis e
5	120	18.2	660	VST2_HEVBU	P29326 hepatitis e
6	109	16.5	123	VST1_HEVME	Q03499 hepatitis e
7	61	9.3	123	VST1_HEVBU	P29325 hepatitis e
8	59	9.0	123	VST1_HEVMY	Q04612 hepatitis e
9	9	1.4	325	ACCA_BACSU	Q34847 bacillus su
10	8	1.2	121	RL7_CLOPE	P06498 human adeno
11	8	1.2	131	E315_ADE05	P28220 sus scrofa
12	8	1.2	153	SORR_PIG	P28279 herpes simp
13	8	1.2	233	UL03_HSV2H	P51249 porphyra pu
14	8	1.2	241	RR2_FORPU	P29786 aedes aegypt
15	8	1.2	254	TRY3_AEDAE	Q09952 caenorhabdi
16	8	1.2	267	YSR4_CAEEL	Q9fuz2 arabidopsis
17	8	1.2	273	DEFC_ARATH	P46316 anthranilic
18	8	1.2	322	ACCA_ANTSP	

c	19	8	1.2	324	1	ACCA_CYACA	O19903 cyanidium c
c	20	8	1.2	324	1	ACCA_PORPU	P51371 porphyra pu
c	21	8	1.2	326	1	ACCA_SYNY3	P74638 synechocyst
c	22	8	1.2	327	1	ACCA_SYNP7	Q54766 synechococc
c	23	8	1.2	331	1	MANI_MOUSE	Q9wu40 mus musculus
c	24	8	1.2	386	1	DGT1_ZYMMO	Q9gm70 oryctolagus
c	25	8	1.2	397	1	S17A_RABIT	Q9rn55 zymomonas m
c	26	8	1.2	414	1	K193_MOUSE	Q9czc8 mus musculus
c	27	8	1.2	448	1	PDFT_CANAL	P78589 candida alb
c	28	8	1.2	449	1	CAPT_PSEPU	P28269 pseudomonas
c	29	8	1.2	459	1	F3A2_STRPU	Q04073 strongyloce
c	30	8	1.2	474	1	RCA_ARATH	P10896 arabidopsis
c	31	8	1.2	505	1	EX7L_CAUCR	Q9a649 caulobacter
c	32	8	1.2	507	1	MKR3_HUMAN	Q33064 homo sapien
c	33	8	1.2	572	1	PDFT_USUMA	Q92459 ustilago ma
c	34	8	1.2	584	1	ACES_RABIT	Q29499 oryctolagus
c	35	8	1.2	614	1	ACES_HUMAN	P22303 homo sapien
c	36	8	1.2	614	1	ACES_RAT	P37136 rattus norv
c	37	8	1.2	732	1	PSAB_CYAME	Q85fy6 cyanidiosch
c	38	8	1.2	735	1	DUR3_YEAST	P33413 saccharomyc
c	39	8	1.2	800	1	T2D4_HUMAN	Q15542 homo sapien
c	40	8	1.2	928	1	CHS2_EXODE	P30601 exophiala d
c	41	8	1.2	1040	1	YEGN_ECOLI	P76398 escherichia
c	42	8	1.2	1625	1	CTPI_MYCTU	Q10900 mycobacteri
c	43	8	1.2	1783	1	RAA3_CHLRE	Q9fec4 chlamydomon
c	44	8	1.2	3969	1	HRX_HUMAN	Q03164 homo sapien
c	45	7	1.1	26	1	MEL_APIDO	P01502 apis dorsat

ALIGNMENTS

RESULT 1

VST2_HEVME VST2_HEVME STANDARD; PRT; 659 AA.

AC Q03500;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor.
OS Hepatitis E virus (strain Mexico) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31768;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93079857; PubMed=1448913;
RA Huang C.C., Nguyen D., Fernandez J., Yun K.Y., Fry K.E.,
RA Bradley D.W., Tam A.W., Reyes G.R.;
RT "Molecular cloning and sequencing of the Mexico isolate of hepatitis
E virus (HEV).";
RL Virology 191:550-558(1992).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M74506; AAA45732.1; -
DR PIR; B44212; B44212.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW Signal.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 659 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 659 AA; 70640 MW; CF75E75EFD8FBE2C CRC64;


```
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Pakistan) (HEV).
OS Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=33774;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.J., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID. CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M80581; AAA45727.1; -.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW Signal.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70980 MW; 8085BC53CFB46FD3 CRC64;

Alignment Scores:
Pred. No.: 2.93e-139 Length: 660
Score: 149.00 Matches: 149
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.61% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST2_HEVPA (1-660)
Qy 490 CTGCGCTGACGAGCGTACTACTACATTAATGCGCACAGAGCGCTCCAATTATGCA 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183
Qy 550 CAGTACCGGGTGGCCGGCTACTATCGTTACCGGCCCTAGTGCCTAATGCGATTGGA 609
Db 184 GlnTyrArgValAlaAlaGalaThrIleArgTyrArgProLeuValProAsnAlaValGly 203
Qy 610 GGCTATGCTATATCAATTCCTCTGGGCTCAACACCAACCCCTACATCTGTTGAC 669
Db 204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
Qy 670 ATGAATTCATTAATTCATGATGTCAGGATTCCTGTTCACCTGGCATGACATCTGAA 729
Db 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
Qy 730 TTGTCATCCCAAGGAGCGCTTCACTACCGCATCAAGTTGGCGCTCGGTGAGACA 789
Db 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 263
Qy 790 TCTGTGTGTCTGAGGAGAGACCACTCCGCTCTTGTCATGTTATGCATCATGGCTCT 849
Db 264 SerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283
Qy 850 CCAGTTAACTCCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGCACTTTGCC 909
Db 284 ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla 303
```

```
Qy 910 TTAGAGCTTCAGTTTCGCAATCTCACC 936
Db 304 LeuGluLeuGluPheArgAsnLeuThr 312
RESULT 3
VST2_HEVRH
ID VST2_HEVRH STANDARD; PRT; 485 AA.
AC 000270;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Structural protein 2 (Fragment).
OS Hepatitis E virus (isolate Rhesus) (HEV).
OS Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31766;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K.,
RA Min K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90274; BAA20910.1; -.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
FT NON_TER 1 1
FT NON_TER 485 485
SQ SEQUENCE 485 AA; 52317 MW; 5A4FD9273AC74F94 CRC64;

Alignment Scores:
Pred. No.: 2.14e-114 Length: 485
Score: 124.00 Matches: 124
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.82% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST2_HEVRH (1-485)
Qy 565 GCGCTACTATCCGTTACCGGCCCTAGTGCCTAATGCGATTGAGGCTATGCTATATCC 624
Db 57 ArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSer 76
Qy 625 ATTTCTTCTGGCTCAACACCAACCCCTACATCTGTGACATGAATTCCTACTACT 684
Db 77 IleSerPheTrpProGlnThrThrThrProThrProThrSerValAspMetAsnSerIleThr 96
Qy 685 TCCACTCATGTGAGGATTCCTTGTCAACCTGGCATAGCATCTGAATTTGGTCCCAAGC 744
Db 97 SerThrAspValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSer 116
Qy 745 GAGCGCTTCACTACCGCAATCAAGTTGGCGCTCGGTGAGACATCTGGTGTGCTGAG 804
Db 117 GluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGlu 136
Qy 805 GAGGAAGCCACCTCCGCTCTTGTTCATGTTATGCATACATGCTCTCCAGTTAACTCTAT 864
Db 137 GluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyr 156
Qy 865 ACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGGACTTTGCTTAGAGCTTGAGTTT 924
```

```
|||||
Db 157 ThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPhe 176
QY 925 CGCAATCTCACC 936
Db 177 ArgAsnLeuThr 180
RESULT 4
VST2_HEVMY STANDARD; PRT; 660 AA.
AC Q04611;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE Structural protein 2 precursor (ORF2)
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31769;
RN [1]
RP MEDLINE=93227573; PubMed=8470371;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikihisa T., Winn K.;
RT "Sequence and gene structure of the hepatitis E virus isolated from
RT Myanmar.";
RL Virus Genes 7:95-109(1993).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; D10330; BAA01174.1; -.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW SIGNAL.
FT CHAIN 1 22 BY SIMILARITY.
FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70998 MW; 3AB2A4EA255C6253 CRC64;
Alignment Scores:
Pred. No.: 2,05e-114 Length: 660
Score: 124.00 Matches: 124
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.82% Indels: 0
DB: 1 Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST2_HEVMY (1-660)
QY 565 CGCGTACTATCCGTACCGGCCCTAGTCCCTAATGAGTGGAGGCTATGCTATATCC 624
Db 189 ArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSer 208
QY 625 ATTTCCTTCTGGCCCTCAACACACACACCCCTAGATCTGTGTGACATGAATCCATTACT 684
Db 209 IleSerPheTyrProGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThr 228
QY 685 TCACATGATGTCAGATCTTCTTCAACTGGCATAGCATCTGAATGTGTCATCCCAAGC 744
Db 229 SerThrAspValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSer 248
QY 745 GAGCGCTTCACTACCGCAATCAAGTTGGCGCTGGTGGACATCTGGTCTGCTGAG 804
Db 249 GluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGlu 268
```

```
QY 805 GAGGAAGCCACCTCCGGTCTCTGTCATGTTATGCATACATGGCTCTCCAGTTAACTCCTAT 864
Db 269 GluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyr 288
QY 865 ACCAATACCCCTTATATACCGGTGCCCTTGGCTTACTGGACTTTGCCCTTAGAGCTTAGAGTTT 924
Db 289 ThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPhe 308
QY 925 CGCAATCTCACC 936
Db 309 ArgAsnLeuThr 312
RESULT 5
VST2_HEVBV STANDARD; PRT; 660 AA.
AC F29326;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE Structural protein 2 precursor (ORF2)
OS Hepatitis E virus (strain Burma) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31767;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; M73218; AAA45736.1; -.
DR FIR; C40778; VHWNH2.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW SIGNAL.
FT CHAIN 1 19 POTENTIAL.
FT CHAIN 20 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70978 MW; 5832A013CCC4A61C CRC64;
```

Alignment Scores:

Pred. No.:	1,94e-110	Length:	660
Score:	120.00	Matches:	120
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	18.21%	Indels:	0
DB:	1	Gaps:	0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST2_HEVBV (1-660)

```
QY 490 CTGCGCTGCGAGGACGGTACTAATACTACATTATGGCCACAGAGCGCTCCCAATTATGCA 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183
QY 550 CAGTACCGGTTGCCCGCGCTACTATCCGTTACCGGCCCTAGTGCCTAATCAGTTGGA 609
Db 184 GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203
```

```
QY 610 GGCTATGCTATATCATTTCTTCTGCGCTCAACCAACACCCCTACATCTGTTGAC 669
Db |||||||SerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCCTACTTCCACTGATGTGAGGATCTTGTTCACCTGGCATGACATCTGAA 729
Db |||||||ThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
QY 730 TTGTCTATCCCAAGGAGCGCTTCACTACCGCAATCAAGTTGGCGCTCGGTTGAGACA 789
Db |||||||ProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 263
QY 790 TCTGCTGTCTGAGGAGGAGCACCTCCGCTCTTGTGTCATGTTATGATACATGGCTCT 849
Db |||||||ValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283

RESULT 6
VSTL_HEVME
ID VSTL_HEVME STANDARD; PRT; 123 AA.
AC Q03499, (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE Structural protein 1.
OS Hepatitis E virus (strain Mexico) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31768;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93079857; PubMed=1448913;
RA Huang C.C., Nguyen D., Fernandez J., Yun K.Y., Fry K.E.,
RA Bradley D.W., Tam A.W., Reyes G.R.;
RT "Molecular cloning and sequencing of the Mexico isolate of hepatitis
RT E virus (HEV).";
RL Virology 191:550-558(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74506; AAA45731.1; -.
DR PIR; C44212; C44212.
DR InterPro; IPR003384; HEV ORF2.
DR Pfam; PF02444; HEV ORF2; 1.
SQ SEQUENCE 123 AA; 12714 MW; C888F5D638852A68 CRC64;

Alignment Scores:
Pred. No.: 2,1e-99 Length: 123
Score: 109.00 Matches: 109
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.54% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VSTL_HEVME (1-123)

QY 2 TGGCGCCTAGGCCTCTTTTGTGTTGTTTCTTGTCTGCTATGTTGCGCGGCAC 61
Db |||||||LeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGACGGGTACGCGTCTGGCGCGCTGCTGGCGCGCAGCGCGGTACCGCGGTGTT 121
Db |||||||ArgProValSerArgLeuAlaAlaValValGlyGlyAlaAlaValProAlaValVal 54
QY 122 TCTGGGGTGACGGGTGATCTTCAGCCCTTTCGCAATCCCTATATTCATCCCAACCAACC 181
Db |||||||ValThrGlyLeuIleLeuSerProSerGlnSerProIlePheIleGlnProThr 74
```

```
QY 182 CCTTTGCCACAGAGTTGCGCTGCGTCTGGACCTCGGCTTGCACCAACAGCCC 241
Db |||||||ProLeuProGlnThrLeuProLeuArgProGlyLeuAspLeuAlaPheAlaAsnGlnPro 94
QY 242 GGCACACTTGCTCCACTTGGCGAGATCAGGCCAGCCCTCCGCTGCTCCGCTGCC 301
Db |||||||HisLeuAlaProLeuGlyGluIleArgProSerAlaProLeuProProValAla 114
QY 302 GACCTGCCACAGCCGGGCTGCGGCGC 328
Db |||||||AspLeuProGlnProGlyLeuArgArg 123

RESULT 7
VSTL_HEVBU
ID VSTL_HEVBU STANDARD; PRT; 123 AA.
AC P29325;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 1 (ORF3).
OS Hepatitis E virus (strain Burma) (HEV), and
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31767, 33774;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Burma;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Pakistan;
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M73218; AAA45735.1; -.
DR EMBL; M80581; AAA45726.1; -.
DR PIR; B40778; VHWHE.
DR InterPro; IPR003384; HEV ORF2.
DR Pfam; PF02444; HEV ORF2; 1.
SQ SEQUENCE 123 AA; 12676 MW; 8A5A798B1B74EDES CRC64;

Alignment Scores:
Pred. No.: 1,06e-51 Length: 123
Score: 61.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.26% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VSTL_HEVBU (1-123)

QY 2 TGGCGCCTAGGCCTCTTTTGTGTTGTTTCTTGTCTGCTATGTTGCGCGGCAC 61
Db |||||||CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGACGGGTACGCGTCTGGCGCGCTGCTGGCGCGCAGCGCGGTACCGCGGTGTT 121
```

```
Db 35 ArgProValSerArgLeuAlaAlaValValGlyGlyAlaAlaValProAlaValVal 54
122 TCTGGGGTACCGGGTGGATTCCTCAGCCCTTCGCAATCCCTATATTATTCACCAACC 181
Db 55 SerGlyValThrGlyLeuLeuSerProSerGlnSerProIlePheIleGlnProThr 74
182 CCT 184
75 Pro 75
RESULT 8
VST1_HEVMY
ID VST1_HEVMY STANDARD; PRT; 123 AA.
AC Q04612;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Structural protein 1 (ORF3).
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis B-like viruses.
OX NCBI_TaxID=31769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227573; PubMed=8470371;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikihisa T., Winn K.;
RT "Sequence and gene structure of the hepatitis E virus isolated from
RT Myanmar.";
RL Virus Genes 7:95-109 (1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10330; BAA01173.1; -
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
SQ SEQUENCE 123 AA; 12751 MW; FB81143F0B31F8A4 CRC64;
Alignment Scores:
Pred. No.: 1,03e-49 Length: 123
Score: 59.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.95% Indels: 0
DB: 1 Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST1_HEVMY (1-123)
QY 8 CTAGCCCTTTTGTGTTGTTCTCTGTTCTGCTATGTTGCGCGCCGACGACG 67
Db 17 LeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHisArgPro 36
QY 68 GTCAGCCCTCTGGCGCGCTGTCGGCGCGGACGCGGTACCGCGGTGTTCTGGG 127
Db 37 ValSerArgLeuAlaAlaValValGlyGlyAlaAlaValProAlaValSerGly 56
QY 128 GTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTTCATCCACCAACCCT 184
Db 57 ValThrGlyLeuLeuSerProSerGlnSerProIlePheIleGlnProThrPro 75
RESULT 9
ACCA_BACSU
ID ACCA_BACSU STANDARD; PRT; 325 AA.
AC O34847;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
```

```
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha
DE (EC 6.4.1.2)
GN ACCA OR BS029200.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
RT in the 200 kb rmb-dnaB region.";
RL Microbiology 143:3431-3441 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Fezzari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerz-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinola S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maul C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nöback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256 (1997).
CC -!- FUNCTION: This protein is a component of the acetyl coenzyme A
CC carboxylase complex; first, biotin carboxylase catalyzes the
CC carboxylation of the carrier protein and then the transcarboxylase
CC transfers the carboxyl group to form malonyl-CoA (by similarity).
CC -!- CATALYTIC ACTIVITY: Carboxybiotin carboxyl carrier protein +
CC acetyl-CoA = biotin carboxyl carrier protein + malonyl-CoA.
CC -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC -!- SUBUNIT: ACETYL-CoA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX (BY SIMILARITY).
CC -!- SIMILARITY: TO THE C-TERMINUS OF MAMMALIAN PROTEINYL-CoA
CC CARBOXYLASE BETA CHAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF008220; AAC00341.1; -
DR EMBL; Z99118; CAB14880.1; -
```

DR PIR; G69580; G69580.
DR Subtilisin; BG12557; accA.
DR InterPro; IPR001095; Ac-CoA_carboxylA.
DR Pfam; PF03255; ACCA; 1.
DR PRINTS; PR01069; ACCTRFRASEA.
DR TIGRFAMS; TIGR00513; accA; 1.
KW Fatty acid biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 325 AA; 36333 MW; 9B177DEBA45B5864 CRC64;

Alignment Scores:
Pred. No.: 4.42 Length: 325
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x ACCA_BACSU (1-325)
QY 294 GAGGACGAGCGGCGGCTGGGCTG 268
DB 199 GlyGlySerGlyGlyAlaLeuGlyLeu 207
|||||

RESULT 10
RL7_CLOPE
ID _RL7_CLOPE STANDARD; PRT; 121 AA.
AC Q8XK7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12.
GN RPLL OR CPE2414.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -!- FUNCTION: Seems to be the binding site for several of the factors
CC involved in protein synthesis and appears to be essential for
CC accurate translation (by similarity).
CC -!- SIMILARITY: Belongs to the L12P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP003194; BAB82120.1; -
CC HAMAP; MF 00368; -; 1.
DR InterPro; IPR008932; Ribos L12/7 olig.
DR InterPro; IPR000206; Ribosomal L12.
DR Pfam; PF00542; Ribosomal L12; 1.
DR ProDom; PD001326; Ribosomal_L12; 1.
DR TIGRFAMS; TIGR00855; L12; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 121 AA; 12535 MW; 5AF9F83D2C94AC3E CRC64;

Alignment Scores:
Pred. No.: 50.1 Length: 121
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.21% Indels: 0
DB: Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x RL7_CLOPE (1-121)
QY 83 GCCGTCTGCGCGCGCGCGCGG 106
DB 38 AlaValValGlyGlyAlaAlaAla 45
|||||

RESULT 11
E315_ADE05
ID _E315_ADE05 STANDARD; PRT; 132 AA.
AC P06498;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Early E3B 14.6 kDa protein precursor.
OS Human adenovirus type 5.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
NCBI_TaxID=28285;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85092388; PubMed=2981456;
RA Cladaras C., Wold W.S.M.;
RT "DNA sequence of the early E3 transcription unit of adenovirus 5.";
RL Virology 140:28-43(1985).
RN [2]
RP IDENTIFICATION OF PROTEIN
RX MEDLINE=90177214; PubMed=2309441;
RA Tollefson A.E., Krajcsi P., Pursley M.H., Gooding L.R., Wold W.S.M.;
RT "A 14,500 MW protein is coded by region E3 of group C human
RT adenoviruses.";
RL Virology 175:19-29(1990).
RN [3]
RP SIGNAL SEQUENCE CLEAVAGE SITE.
RX MEDLINE=92148953; PubMed=1531370;
RA Krajcsi P., Tollefson A.E., Anderson C.W., Wold W.S.M.;
RT "The adenovirus E3 14.5-kilodalton protein, which is required for
RT down-regulation of the epidermal growth factor receptor and
RT prevention of tumor necrosis factor cytotoxicity, is an integral
RT membrane protein oriented with its C terminus in the cytoplasm.";
RL J. Virol. 66:1665-1673(1992).
RN [4]
RP PHOSPHORYLATION.
RX MEDLINE=92188523; PubMed=1532104;
RA Krajcsi P., Wold W.S.M.;
RT "The adenovirus E3-14.5K protein which is required for prevention of
RT TNF cytotoxicity and for down-regulation of the EGF receptor contains
RT phosphoserine.";
RL Virology 187:492-498(1992).
RN [5]
RP O-GLYCOSYLATION.
RX MEDLINE=92263760; PubMed=1533979;
RA Krajcsi P., Tollefson A.E., Wold W.S.M.;
RT "The E3-14.5K integral membrane protein of adenovirus that is
RT required for down-regulation of the EGF receptor and for prevention
RT of TNF cytotoxicity is O-glycosylated but not N-glycosylated.";
RL Virology 188:570-579(1992).
RN [6]
RP COMPLETE GENOME.
RX MEDLINE=92087470; PubMed=1727603;
RA Chroboczek J., Bieber F., Jacrot B.;
RT "The sequence of the genome of adenovirus type 5 and its comparison
RT with the genome of adenovirus type 2.";
RL Virology 186:280-285(1992).
CC -!- FUNCTION: Down-regulates the EGF receptor and prevents cytotoxicity
CC by TNF.
CC -!- PTM: Phosphorylated on serine; O-glycosylated, but not N-
CC glycosylated.
CC -!- SIMILARITY: TO EQUIVALENT PROTEIN IN OTHER ADENOVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; M73260; -; NOT ANNOTATED_CDS.
DR EMBL; X03002; CAA26786.1; -
DR InterPro: IPR008131; Adeno_E3_14_5.
DR Pfam; PF04834; Adeno_E3_14_5; 1.
KW Early protein; Transmembrane; Glycoprotein; Phosphorylation; Signal.
FT SIGNAL 1 18
FT CHAIN 19 132 EARLY E3B 14.6 kDa PROTEIN.
FT TRANSMEM 51 75 POTENTIAL.
SQ SEQUENCE 132 AA; 14750 MW; 524690C4AD9B9A74 CRC64;

Alignment Scores:
Pred. No.: 49.5 Length: 132
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x E315_ADE05 (1-132)

QY 1097 CCGCGCGGACCTTCACCTACCCCA 1074

Db 112 ProArgProThrSerProthrPro 119

RESULT 12
SORB_PIG
ID SORB_PIG STANDARD; PRT; 153 AA.
AC P28220:
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sorbin.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RC TISSUE=Intestine;
RX MEDLINE=92007891; PubMed=1915377;
RA Vagne-Descroix M., Pansu D., Joernvall H., Carlquist M., Guignard H.,
RA Jourdan G., Desvigne A., Collinet M., Callet C., Mutt V.,
RT "Isolation and characterisation of porcine sorbin."
RL Eur. J. Biochem. 201:53-59(1991).
CC -!- FUNCTION: Increases water and sodium absorption in the intestine
CC and gall-bladder.
CC -!- SIMILARITY: Contains 1 SoHo domain.
DR PIR; S17837; S17837.
DR InterPro; IPR003127; Sorb.
DR Pfam; PF02208; Sorb; 1.
DR ProDom; PD016158; Sorb; 1.
DR SMART; SM00459; Sorb; 1.
DR PROSITE; PS50831; SOHO; 1.
FT DOMAIN 1 46 SOHO.
SQ SEQUENCE 153 AA; 17481 MW; A49C219455BB1417 CRC64;

Alignment Scores:
Pred. No.: 48.5 Length: 153
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x SORB_PIG (1-153)

QY 119 CCACCGCGGTACCGCGTGCGC 96

Db 93 ProProValProLeuArg 100

RESULT 13
UL03_HSV2H

ID UL03_HSV2H STANDARD; PRT; 233 AA.
AC P28279;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Protein UL3.
GN UL3.
OS Herpes simplex virus (type 2 / strain HG52).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113549; PubMed=1662697;
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;
RT "Comparative sequence analysis of the long repeat regions and
RT adjoining parts of the long unique regions in the genomes of herpes
RT simplex viruses types 1 and 2."
RL J. Gen. Virol. 72:3057-3075(1991).
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL3,
CC HSV-2 UL3, EBV-1 60, AND VZV 58.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; D10470; BAA01266.1; -
DR PIR; JQ1496; WMBEHK.
DR InterPro; IPR005035; Herpes UL3.
DR Pfam; PF03369; Herpes UL3; 1.
SQ SEQUENCE 233 AA; 25649 MW; B38F613BD839AA24 CRC64;

Alignment Scores:
Pred. No.: 45.7 Length: 233
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x UL03_HSV2H (1-233)

QY 421 TCACCCCTGACATCCTCTGTGCC 444

Db 75 SerProLeuThrSerSerValAla 82

RESULT 14

ID RR2_PORPU STANDARD; PRT; 241 AA.
AC P51249;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S2.
GN RPS2.
OS Porphyra purpurea.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Avonport;
RA Reith M.E., Munnholland J.;

RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast

RT genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -!- SIMILARITY: Belongs to the S2P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U36804; AAC08135.1; --
DR PIR; S73170; S73170.
DR HAMAP; MF_00291; --; 1.
DR InterPro; IPR001865; Ribosomal_S2.
DR InterPro; IPR005706; Ribosomal_S2_b/o.
DR Pfam; PF00318; Ribosomal_S2; 1.
DR PRINTS; PR00395; RIBOSOMALS2.
DR TIGRFAMs; TIGR01011; rpsB_bact; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Ribosomal protein; Chloroplast.
SQ SEQUENCE 241 AA; 27198 MW; FA17DA9C352FC866 CRC64;

Alignment Scores:
Pred. No.: 45.5 Length: 241
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x RR2_PORPU (1-241)

QY 1147 CCGACAGATAATTCGTCGGCT 1170
Db 231 ProThrGluLeuIleSerSerAla 238

RESULT 15
TRY3_AEDAE
ID TRY3_AEDAE STANDARD; PRT; 254 AA.
AC P29786;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Trypsin 3A1 precursor (EC 3.4.21.4).
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
OX NCBI_TaxID=71159;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97242546; PubMed=9087545;
RA Kalhok S., Tabak L.M., Prosser D.E., Brook W., Downer A.E.R.,
RA White B.N.;
RT "Isolation, sequencing and characterization of two cDNA clones coding
RT for trypsin-like enzymes from the midgut of Aedes aegypti.";
RL Insect Mol. Biol. 2:71-79(1993).
CC -!- FUNCTION: Major function may be to aid in digestion of the blood
CC meal.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: Midgut.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; X64362; CAA45714.1; --
DR PIR; S19890; TRWV3Y.
DR HSPSP; P00763; IDPO.
DR MEROPS; S01.112; --
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal; Zymogen; Digestion;
KW Multigene family.
FT SIGNAL 1 20 POTENTIAL.
FT PROPEP 21 27 ACTIVATION PEPTIDE.
FT CHAIN 28 254 TRYPSIN_3A1.
FT ACT_SITE 68 68 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 113 113 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 209 209 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 53 69 BY SIMILARITY.
FT DISULFID 178 194 BY SIMILARITY.
FT DISULFID 205 229 BY SIMILARITY.
FT SITE 203 203 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 254 AA; 26939 MW; 43AAF13B8DE7D485 CRC64;

Alignment Scores:
Pred. No.: 45.1 Length: 254
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x TRY3_AEDAE (1-254)

QY 1767 CGTCCGGTCCCATTTCTCGGC 1790
Db 47 ArgSerGlyArgHisPheCysGly 54

Search completed: August 16, 2004, 13:48:51
Job time : 26.2362 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:41:04 ; Search time 19.9849 Seconds
(without alignments)
10229.686 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096

Perfect score: 659

Sequence: 1 ATGGCGCCTAGGCTCTTTT.....GTAAACTCGGAGCTTGAG 1980

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Word size: 1

Total number of hits satisfying chosen parameters: 663654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p_model -DEV=xlp
-Q=/cgn2_1/USPRO_pool_p/US09851410/runat_16082004_125718_9682/app_query.fasta_1.4238
-DB=Issued Patents AA -QMT=fastan -SUFFIX=oli.rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09851410@cgn1_1_45@runat_16082004_125718_9682 -NCPUL=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pap:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	659	100.0	659	1	US-08-240-049B-16
2	659	100.0	659	1	US-08-259-148A-20
3	659	100.0	659	1	US-08-484-054-20
4	659	100.0	659	2	US-07-876-941A-20
5	659	100.0	659	3	US-08-477-292-14
6	659	100.0	659	4	US-07-870-985A-20
7	548	83.2	548	3	US-08-477-292-16
8	512	77.7	660	3	US-08-542-634-14
9	512	77.7	660	5	PCT-US95-13703-14
10	438	66.5	660	4	US-09-462-606-48
11	435	66.0	435	1	US-08-259-148A-18
12	435	66.0	435	1	US-08-484-054-18

13	435	66.0	435	2	US-07-876-941A-18	Sequence 18, Appl
14	435	66.0	435	4	US-07-870-985A-18	Sequence 18, Appl
15	401	60.8	549	3	US-08-542-634-16	Sequence 16, Appl
16	401	60.8	549	5	PCT-US95-13703-16	Sequence 16, Appl
17	392	59.5	540	3	US-08-542-634-26	Sequence 26, Appl
18	392	59.5	540	5	PCT-US95-13703-26	Sequence 26, Appl
19	377	57.2	525	3	US-08-542-634-28	Sequence 28, Appl
20	377	57.2	525	5	PCT-US95-13703-28	Sequence 28, Appl
21	327	49.6	327	1	US-08-240-049B-14	Sequence 14, Appl
22	327	49.6	327	1	US-08-259-148A-16	Sequence 16, Appl
23	327	49.6	327	1	US-08-484-054-16	Sequence 16, Appl
24	327	49.6	327	2	US-07-876-941A-16	Sequence 16, Appl
25	327	49.6	327	3	US-08-542-634-18	Sequence 18, Appl
26	327	49.6	327	3	US-08-477-292-18	Sequence 18, Appl
27	327	49.6	327	4	US-07-870-985A-16	Sequence 16, Appl
28	327	49.6	327	5	PCT-US95-13703-18	Sequence 18, Appl
29	149	22.6	552	4	US-09-172-699-20	Sequence 20, Appl
30	149	22.6	561	4	US-09-172-699-20	Sequence 20, Appl
31	149	22.6	660	4	US-09-462-606-12	Sequence 12, Appl
32	149	22.6	660	4	US-09-462-606-49	Sequence 49, Appl
33	149	22.6	660	4	US-09-462-606-51	Sequence 51, Appl
34	149	22.6	660	4	US-09-172-699-2	Sequence 2, Appl
35	124	18.8	660	4	US-09-462-606-55	Sequence 55, Appl
36	120	18.2	525	3	US-08-542-634-27	Sequence 27, Appl
37	120	18.2	525	5	PCT-US95-13703-27	Sequence 27, Appl
38	120	18.2	540	3	US-08-542-634-25	Sequence 25, Appl
39	120	18.2	540	5	PCT-US95-13703-25	Sequence 25, Appl
40	120	18.2	549	3	US-08-542-634-15	Sequence 15, Appl
41	120	18.2	549	3	US-08-477-292-15	Sequence 15, Appl
42	120	18.2	549	5	PCT-US95-13703-15	Sequence 15, Appl
43	120	18.2	660	1	US-08-240-049B-15	Sequence 15, Appl
44	120	18.2	660	1	US-08-259-148A-19	Sequence 19, Appl
45	120	18.2	660	1	US-08-484-054-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1

US-08-240-049B-16
; Sequence 16, Application US/08240049B
; Patent No. 5686239
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Tam, Albert W.
; APPLICANT: Yarbough, Patricia O.
; TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,049B
; FILING DATE: 09-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles K. Sholtz
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.35; G32P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico Strain)
; INDIVIDUAL ISOLATE: ORF-2
US-08-240-049B-16

Alignment Scores:
Pred. No.:      0      Length:      659
Score:          659.00  Matches:      659
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      100.00%  Indels:      0
DB:                1      Gaps:        0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-08-240-049B-16 (1-659)

QY 1 ATGCGCCCTAGGCCTCTTTTGTGTTGTTCTCTTGTGTTCTGCTATGTCGCGGCCA 60
Db 1 MetArgProArgProLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20

QY 61 CGACCGGTGACGGGTCCTGCGCGCCCGTCTGCGCGCGCGGCGGTACCGGCGGTGT 120
Db 21 ProThrGlyGlnProSerGlyArgArgGlyArgSerGlyGlyThrGlyGlyGly 40

QY 121 TTCTGGGTGACGGGTTGATTCTCAGCCCTTCGCNAATCCCTATATTCATCAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60

QY 181 CCCTTTGCCCCAGAGCTTCGGCTGCTCGGCTCTGGACCTCGCTCCCAACAGCC 240
Db 61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80

QY 241 CGGCCACTTGGCTCCACTTGGCGAGATCAGGCCCGCCCGCTCGCTCCCGTCCG 300
Db 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArg 100

QY 301 CGACCTGCCACAGCGGGCTGCGGCTGCGAGCTGCGGCTGCGGCTGCGGCTGCGGCTCA 360
Db 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120

QY 361 CCGCTCCCGGACGTTGATTCTCGCGTGCAATTCTACGCGCGCAGTAAATTGTTCTACT 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140

QY 421 TCACCCCTGACATCCTCTGTGCGCTCTGGCCTAAATTTAGTCTCTGTATGACGCCCTT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160

QY 481 AATCCGCTCTGCGCTGACAGCGTACTAATCTACATTCATGCGCCACAGAGGCTCC 540
Db 161 AsnProProLeuLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180

QY 541 AATTATGACAGTACCGGTTGCGCGCTACTATCTCCGTTACCGGCCCTAGTGCCTAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200

QY 601 GCAGTTGGAGGCTATGCTATATCCATTCTTTCTGCGCTTCAACCAACCAACCCCTACA 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220

QY 661 TCTGTTGATGATGAATCCATTACTTCCACTGATGTCAGGATCTTGTTCAACTGGCATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240

QY 721 GCATCTGAATTGGTTCATCCACAGCGCGCTTCTACTACCGCAATCAAGGTTGGGCTCG 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260

QY 781 GTTAGACATCTGGTGTCTGAGGAGGAGGCACTCCGGTCTTGTGTCATGTTATGATA 840
Db 821 ThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGln 640
```


Db 1 MetArgProArgProLeuLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY 61 CCGACCGGTGACGGTGTGCGCGCGTGTGGCGCGCGCGGTACCGCGGTGGT 120
Db 21 ProThrGlyGlnProSerGlyArgArgGlyArgSerGlyGlyThrGlyGlyGly 40
QY 121 TTCGGGGTGCAGGGTGTGATCTCAGCCCTTCCCAATCCCTATATTCACCAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCCTTTCGCCAGACGGTGTGCGTCCGGTCTGGACCTCGCGTTCGCCAACACCGCC 240
Db 61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY 241 CGGCCACTTGGCTCCACTTGGCGAGATCAGGCCCGCGCGCTCGCTCGCTCCCGTCCG 300
Db 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
QY 301 CGACCTGCCAGCGGGCTGCGCGCTGACGGCTGACGGCTGCGCGCTGCCCATGACACCTCA 360
Db 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCCGTCCGGACGTGTGATCTCGCGGTGCAATTTCTAGCGCGCCAGTATAAATTTGTCTACT 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCACCCCTGACATCCTCTGTGGCTCTGGCACTAATTTAGTCGTGTATGACGCCCCCTT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AATCCGCTCTGCGCTGCAGGAGGTACTCAATCTCACATTATGGCCACAGAGCCCTCC 540
Db 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCACAGTACCGGGTTCGCGGTACTATCCGTTACCGCGCCCTAGTGCCTAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCAGTGGAGCTATGCTATATCCATTCTTTCTGGCTCAAAACACACACCCCTACA 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCTGTTGACATGAATTCATCTACTTCCACTGATCTCAGGATCTTGTTCACCTGGCATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCATCTGAATTTGGTCATCCCAAGCGAGCGCTTCACTACCGCAATCAAGTTGGCGCTCG 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTTGAGACATCTGTGTGTCTGAGAGAGAACCCCTCCGCTCTTGCTCATGTTATGCGATA 840
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCCAGTAACTCTATACCAATACCCCTTATACGGTTCGCTGGCTTACTG 900
Db 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTTGCTTAGAGCTTGAGTTTCGCAATCTCCACCATCTGAACCAATACACAGTGTG 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTACTCCAGACATCTGCTGTCATCCCGCGAGGGCGGCGGAGTGCAGGTG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCCAACTGCAGCCACAGGTTTCATGAAGATCTCCACTTTACCGCTTAATGGGGTA 1080
Db 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GGTGAAGTCGCGCGGATAGTCTCAATTTACTTAACCTTGTGTCACAGCTCCTCCGCG 1140

Db 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGCTCCGCGACAGATTAATTTCTGCTCGCTGGCGGCAACTGTTTATTCGCCGCCGCTT 1200
Db 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GTCTCAGCCAATGGCAGCAACCGTGAAGCTCTATACATCAGTGGAGAATGCTCAGCAG 1260
Db 401 ValSerAlaAsnGlyGluProThrValValLeuLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 GATAAGGGTGTGCTATCCCCACGATATCGAATCTTGGTGAATTCGCTGTGTGCTCATTCAG 1320
Db 421 AsplysglyValAlaIleProHisAspIleAspLeuGlyAspSerArgValValIleGln 440
QY 1321 GATTATCACACACAGCATGAGCAGATCGCCGCCACCCGCTCGCTCGCCCATCTCGCCCT 1380
Db 441 AspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
QY 1381 TTTTCTGTTCTCCGAGCAATGATGTACTTGGTGTCTCCCTCACTGCGACCGAGTATGAC 1440
Db 461 PheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAsp 480
QY 1441 CAGTCCACTTACGGGTCTCAACTGCGCCCGTATATCTCGGACAGCGTGTGGTG 1500
Db 481 GlnSerThrTyrGlySerSerThrThrGlyProValTyrIleSerAspSerValThrLeuVal 500
QY 1501 AATGTTCCGACTGCGCGCGCGAGCGTAGCCGATGCTGCTGCTCCAAAGTCAACCTC 1560
Db 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeu 520
QY 1561 GACGGCGCGCCCTCCGACTGTTTGACAATATTCCAAGACATTTCTTGTGCTCCCTT 1620
Db 521 AspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuProLeu 540
QY 1621 CGTGGCAAGCTCTCCTTTTGGGAGCGCGCACAAACAAAGCAGGTTATCTTATAATAT 1680
Db 541 ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyr 560
QY 1681 AATFACTGCTAGTAGCAGGATTTCTGATTGAAAATGCTCGCGGCCATCGGGTCCCAT 1740
Db 561 AsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIle 580
QY 1741 TCAACCTATACACAGCTTGGCGCGCTCGCTCGCTCGCTTCTGCGCGCGCGGTTTG 1800
Db 581 SerThrTyrThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeu 600
QY 1801 GCTCCAGCTCCCGCTGCTGCTCTCGAGGATACTTTTGTATTCGCGGGCGGCGCAC 1860
Db 601 AlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHis 620
QY 1861 ACATTTGATGACTTCTCCCTGAATGCGCGCTTAGCCCTCCAGGTTGTGTTCAG 1920
Db 621 ThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGln 640
QY 1921 TCAACTGCTCGTGAAGTCCAGCGCTTAAAGCTTAAGTGGGTGGAACCTCCGAGGTTG 1977
Db 641 SerThrValAlaGluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 659

RESULT 4

US-07-876-941A-20
; Sequence 20, Application US/07876941A
; Patent No. 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; TITLE OF INVENTION: Antibodies
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/876,941A
FILING DATE: 01-MAY-1992
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.33
TELEPHONE/DOCKET INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 9
US-07-876-941A-20

Alignment Scores:
Pred. No.: 0 Length: 659
Score: 659.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-07-876-941A-20 (1-659)

QY 1 ATGGCCCTAGCCCTCTTTGCTGTGTTCTCTTGTCTGCTTATGTTGCCCGCGCA 60
DB 1 MetArgProArgProLeuLeuLeuLeuPheLeuPheLeuProMetLeuProAlaPro 20
QY 61 CGACCGGTACGCGCTCTGCGCGCGCTCTGTTGGCGCGCGAGCGGCTACCGCGGTGT 120
DB 21 ProThrGlyGlnProSerGlyArgArgArgGlyArgArgSerGlyGlyThrGlyGly 40
QY 121 TTCTGGGTGACCGGGTGAATTTCTCAGCCCTTCGAAATCCCTATATTCATCCAAACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTTGCCACAGCTTCCCGCTCGCTCGCGCTCTGGACCTTCGCGCTTCGCCAACAGCC 240

61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY 241 CGGCCACTTGGCTCCACTTGGCGAGATCAGGCCCGCCCGCTCCGCTCCCTCCGTCGC 300
DB 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArg 100
QY 301 CGACCTGCCACAGCGGGGCTCGCGCGCTGACGGCTGTGGCGCTCCCTCCATCACCTCA 360
DB 101 ArgProAlaThrAlaGlyAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCCGTCGGGACGTTGATTTCTCGCGTGCAATTTCTACGCCGACGATTAATTTGTCTACT 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCACCCCTGACATCTCTGTGGCTCTGSCACTAATTTAGTCTCTGTATGACGCCCCCTT 480
DB 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AATCCGCTCTCCGCTGCGAGACGGTACTAATATCTACATATTGGCCACAGAGGCTCC 540
DB 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGACAGTACGGGTTCCCGCGCTACTATCCGTTACGGGCCCTAGTGCCTAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTCGCGCTCAACAAACACCAACCCCTACA 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrProThr 220
QY 661 TCTGTTGACATGAATTTCCATTTACTTCCACTGATGTGAGGATTTCTTGTCAACCTGGCATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCATCTGAATGCTCATCCAGAGCGCCTTACTACCGCAATCAAGTTGGGCTCG 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTTGAGACATCTGGTGTGCTGAGGAGGAAGCCACTCCGCTTGTGTATGTTATGCATA 840
DB 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGTCCCTTGGCTTACTG 900
DB 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCTTAGACCTTGAGTTTCGCAATCTCACACCTGTAAACACCAATACAGTGTG 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTACTCCAGCACTGCTGTCACCTCCGCGCGAGGGCGCGAGGACTCGGAGCTG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCCAACTGCAGCCACCAGGTTTCATGAAAGATCTCCACTTTACCGGCTTAAATGGGTA 1080
DB 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GGTGAGTCCGCGCGGAGTAGCTTAACATTAATTAACCTTGCTGACACGCTCCCTCGGC 1140
DB 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGCTCCCGACAGAAATTAATTTGCTCGCTGCGCGGCAACTGTTTATTTCCGCGCGGTT 1200
DB 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GTCTCAGCCAATGCGCGCCAAACCGTGAAGCTCTATATCATCAGTGGAGAACTCAGCAG 1260
DB 401 ValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAGGCTGTGCTATCCCGACGATATCCATCTTGTGATTCCGCTGTGCTTCATTCCAG 1320
DB 421 AspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValValIleGln 440

Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
QY 661 TCTGTTGACATGAATTCCTACTTCTCCTGATGTCAGGATCTCTGTTCAACCTGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleuValGlnProGlyIle 240
QY 721 GCATCTGAATTGGTCATCCCAAGCAGCGCCCTTCACTACCGCAATCAAGGTTGGCGCTCG 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTTGAGACATCTGGTGTCTCTGAGGAGGAAGCACCCTCGGTCTGTGTCATGTATGCATA 840
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCTCCAGTTAACTCTATACCAATACCCCTTATACCGGTCCCTCTGGCTTACTG 900
Db 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAACACCAATACAGTGTG 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTACTCCAGCAGCTCTGCTGCTACCTCGGCCGAGGGCGCAGCGGACTCGGAGCTG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCCAACTGCGAGCCACCAAGGTTTCATGAAGATCTCCACTTTACCGCCCTTAATGGGTA 1080
Db 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GTGAAGTCGGCGGGGATAGCTCTACATTACTTAACCTTGCTGACACGCTCTCTCGGC 1140
Db 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGCTCCGACAGAAATTAATTTCTGCTGGCTGGGGGCACTGTTTATTCGGCGCGGTT 1200
Db 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GTCTCAGCCAAATGGCGAGCAACCGTGAAGCTCTATACATCAGTGGAGAACTCTCAGCAG 1260
Db 401 ValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAGGGTGTGCTATCCCCACGATATCGATCTTGTGTGATTCGCGTGTGTCATTCA 1320
Db 421 AspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValIleGln 440
QY 1321 GATTATGACACGACATGACGAGATCGGCCACCCGCTGCGCTCGCCATCTCGGCT 1380
Db 441 AspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
QY 1381 TTTTCTGTTCTCCGAGCAATGATGTACTTTGGCTGTCCCTCACTGCGAGCCGAGTATGAC 1440
Db 461 PheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAsp 480
QY 1441 CAGTCCACTTACGGTGTCAACTGGCGCGGTTTATATCTCGGACAGCGTGTGCTTGGTG 1500
Db 481 GlnSerThrTyrGlySerThrGlyProValTyrIleSerAspSerValThrLeuVal 500
QY 1501 AATGTTGCGACTGGCGCGCAGCGCTAGCCGATCGCTTGACTGCTCCAAAGTCAACCTC 1560
Db 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeu 520
QY 1561 GACGGCGGCGCCCTCCCGACTGTTGAGCAATATCCAGCAATCTTTGTGCTCCCGCTT 1620
Db 521 AspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuProLeu 540
QY 1621 CGTGCAAGCTCTCTTTTGGAGCGCGGCACAAACAAAGCAGGTTATCTTTATTAATAT 1680
Db 541 ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyr 560
QY 1681 AATACTACTGCTAGTACAGATTTCTGATTAAGAAATGCTGCGCGCATCGGGTCCGCATT 1740
Db 561 AsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIle 580

QY 1741 TCAACCTATACCAACAGGCTTGGGGCCGTCGCGTCCGTCATTTCTCGGCGCGGCTTTG 1800
Db 581 SerThrTyrThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeu 600
QY 1801 GCTCCACGCTCCGCCCTCGCTCTGCTGGAGGATACCTTTTGATTTATCCGGGGCGGCGCAC 1860
Db 601 AlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHis 620
QY 1861 ACATTTGATGACTTTCGCCCTGAAATGCCCGCGCTTTAGGCTCCAGGTTGTGCTTTCCAG 1920
Db 621 ThrPheAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGln 640
QY 1921 TCAACTGTCTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGTAAACTCGGAGTTG 1977
Db 641 SerThrValAlaGluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 659

RESULT 6
US-07-870-985A-20
; Sequence 20, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 659 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 9
; US-07-870-985A-20

Alignment Scores:
Pred. NO.: 0 Length: 659
Score: 659.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-07-870-985A-20 (1-659)

QY 1 ATGGCGCCTAGCCCTCTTCTGCTGTTCTCTGTTCTCTGTTCTGCTATGTCGCCGCGCCA 60
Db 1 MetArgProArgProLeuLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY 61 CCGACCGGTGACCGCTCGCGCGCGCTCGTGGCGCGCGAGCGGCGGTACCGGGCGTGGT 120
Db 21 ProThrGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlyThrGlyGlyGly 40
QY 121 TTCGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyIleHisProThrAsn 60
QY 181 CCCTTTGCCCGAGGTTCGCGCTCGCTCGGGCTCGAGCTCGCTCGCTCGCTCGCTCGCTCG 240
Db 61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY 241 CGGCCACTTGGCTCCACATTGGCGAGATCAGGCCAGCCGCCCTCGCTCGCTCGCTCGCTCG 300
Db 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
QY 301 CGACTGCGCACAGCGGGGTGCGCGCTGACGGCTGACGGCTGCGCGCTGCGCGCTGCGCGCT 360
Db 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCCGTCCCGAGGTGATCTCGCGTCAATCTACCGCCGACCTCGCTCGCTCGCTCGCTCGCT 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyArgGlnTyAsnLeuSer 140
QY 421 TCACCCCTGACATCCTCTGTGGCTCTGGCACTAATTTAGTCTGTATGCGACCCCTCTCT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyAlaAlaProLeu 160
QY 481 AATCGCTCTGCGCTCGAGACGGTACTAATCTACATTAATGCGCACAGAGCCCTCC 540
Db 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATATGACAGTACCGGGTTCGCCGCTACTATCCGTATCCGTACCGCCCTAGTGCTAAT 600
Db 181 AsnTyAlaGlnTyArgValAlaAlaArgAlaThrIleArgTyArgProLeuValProAsn 200
QY 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTCTGGCTCAACCAACCAACACCCCTTACA 660
Db 201 AlaValGlyGlyTyAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
QY 661 TCTGTGTACATGAATTCATTTCCACTGATGTGAGGATCTTTGTTCAACCTGGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCATCTGAATTGCTCATCCCAAGCAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyArgAsnGlnGlyTrpArgSer 260
QY 781 GTTGAGACATCTGGTGTGCTGAGGAGGAGGACCTCGCGTCTTGTGATGTTATGTCATA 840
Db 841
```

```
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCTCCAGTTAACTCCTATACCAATACCCCTTATACCGTGCCTTGGCTTACTG 900
Db 281 HisGlySerProValAsnSerTyThrAsnThrProTyThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTTGCTTAGAGCTTGGATTTGCGCAATCTCACCACTCTGTAACCAATACACGTGTG 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTACTCAGACATGTCGTCACTCCCGCCAGGGCGCGAGCTCGGAGCTG 1020
Db 321 SerArgTySerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCCAACTGCGAGCCACCGGTTTCATGAAGATCTCCACTTTACCGGCTTAATGGGTA 1080
Db 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GGTGAAGTCCGCGCGGATAGCTCTAACATTAACCTTGTGTGACACGCTCCTCGGC 1140
Db 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGTCCCGACAGAAATTAATTTGCTGGCTCGCGGCGCAACTGTTTTATCCCGCGGTT 1200
Db 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlnLeuPheTySerArgProVal 400
QY 1201 GTCTCAGCAATGCGAGCCCAACCGTGAAGCTCTATACATCAGTGGAGAATGCTCAGCAG 1260
Db 401 ValSerAlaAsnGlyGluProThrValLysLeuTyThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAAGGGTGTGTGTTATCCCCACGATATCGATCTTGTGTGATTCGCTGTGTGCTCATTCAG 1320
Db 421 AspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValValIleGln 440
QY 1321 GATTATGACAAACAGCATGACAGAGATCGGCCCAACCCGTCGCTCGCGCTCTCGGCT 1380
Db 441 AspTyArgAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
QY 1381 TTTTCTGTCTCCGAGCAATGATGATCTTGGTGTCTCCCTCAGTCAGCCGAGTATGAC 1440
Db 461 PheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGlnTyArg 480
QY 1441 CAGTCCACTTACGGGTCTCAACTGGCCCGGTTTATATCTCGGACAGCGTCACTTGGTG 1500
Db 481 GlnSerThrTyGlySerSerThrGlyProValTyIleSerAspSerValThrLeuVal 500
QY 1501 AATGTTGCGATGCGCGCAGCGCTAGCCGATCGCTTGACTGTGTCCAAAGTCAACCTC 1560
Db 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeu 520
QY 1561 GACGGGCGCCCTCCGACTGTTGAGCAATATTCGAAGACATTTCTTGTGCTCCCTT 1620
Db 521 AspGlyArgProLeuProThrValGlnGlnTySerLysThrPhePheValLeuProLeu 540
QY 1621 CGTGGCAAGCTCTCCTTTTGGAGCGCGGCAACAAAGAGAGGTTATCCTTATATATAT 1680
Db 541 ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyProTyAsnTy 560
QY 1681 AATACTACTGTAGTACGACAGATCTTGATGAAATGTCGCGGCAATCGGCTCGCCATT 1740
Db 561 AsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIle 580
QY 1741 TCAACCTATACACAGGCTTGGGCGCGTCCGTCGCCATTTCTGCGCGCGCGGTTTG 1800
Db 581 SerThrTyThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeu 600
QY 1801 GCTCCAGCCTCGCCCTCGCTCTGCTGAGGATACCTTTTGATTATCCGGGGCGGCGCAC 1860
Db 601 AlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyProGlyArgAlaHis 620
QY 1861 ACATTGTGATCTCTGCGCTGAATGCGCGCTTTAGGCCTCCAGGGTGTGCTTCCAG 1920
Db 621 ThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGln 640
```

QY 1921 TCAACTGTGCTGAGCTCCAGCGCCCTTAAAGTTAAGTGGGTAAAACTCGGGAGTTG 1977
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
641 SerThrValAlaGluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 659

RESULT 7

US-08-477-292-16
; Sequence 16, Application US/08477292
; Patent No. 6291641
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas
; APPLICANT: McAtee, Patrick
; APPLICANT: Yarbough, Patrice
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cathleen M. Desjardins, M.D.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: US
; ZIP: 94063

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,292

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/327,952

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, Allan A

; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: G32P5

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)369-9500
; TELEFAX: (415)368-0709

; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids

; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: protein

; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E virus (Mexico strain)

; INDIVIDUAL ISOLATE: r62kDa, FIGURE 4
US-08-477-292-16

Alignment Scores:

Pred. No.: 0 Length: 548
Score: 548.00 Matches: 548
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.16% Indels: 0
DB: 3 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-08-477-292-16 (1-548)

QY 334 GCTGTGGCGCTGCCATGACACCTCACCGTCCCGAGGTTGATCTCGCGGTGCAATT 393
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 AlaValAlaProAlaHisAspThrSerProValProAspValAspSerArgGlyAlaIle 20

QY 394 CTACGCCGCCAGTATATTGTCTACTTCACCCCTGACATCTCTGTGGCCTCTGGCACT 453

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaIleSerGlyThr 40

QY 454 AATTTAGTCTGTATGACAGCCCCCTTAATCCCGCTCTGCCGTGACGAGCGGTACTAAT 513
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
41 AsnLeuValLeuTyrAlaAlaProLeuAsnProProLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACTCACATTATGGCCACAGAGCGCTCCAATTATGCACAGTACCGGGTTGCCCGCTACT 573
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaAlaArgAlaThr 80
QY 574 ATCCGTTACGGGCCCTAGTGCTTAATGAGTGGAGGCTATGCTATATCCATTCTTTC 633
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100
QY 634 TGGCCTCAAAACAACAACCCCTACATCTGTGTGACATCAATTCCTACTTCTCACTGAT 693
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
101 TrpProGlnThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp 120
QY 694 GTCAGGATTCTGTTCAACTGGCATAGCATCTGAATTCGTATCCCAAGCGAGCGCCTT 753
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTACCGCAATCAAGGTTGGCGCTCGGTTGACACATCTGTTGCTGAGGAGGAGGCC 813
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
141 HisTyrArgAsnGlnGlyTrpArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCCGCTCTTGTTCATGTTATGCATACATGGCTCTCCAGTTAACTTCACTATACCAATACC 873
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
161 ThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThr 180
QY 874 CTTTATACCGGTGCCCTTGCTTACTGGACTTTTGGCTTAGAGCTTAGAGTTGCGAATCTC 933
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200
QY 934 ACCACCTGTAAACACCAATACAGTGTGTCCTGCTACTCCAGCAGCTGCTGCTCACTCCGCC 993
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
201 ThrThrCysAsnThrAsnThrArgValSerArgTyrSerSerThrAlaAlaHisSerAla 220
QY 994 CGAGGGCGCGACGGGACTCGGAGGTGACACCAACTGCAGCCACCAGCGTTTCATGAAGAT 1053
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
221 ArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrAlaArgPheMetLysAsp 240
QY 1054 CTCCTACTTACCGGCTTAATGGGTAGGTGAAGTCGGCGGGGATAGCTCTAAACATTA 1113
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 LeuHisPheThrGlyLeuAsnGlyValGlyGluValGlyArgGlyIleAlaLeuThrLeu 260
QY 1114 CTTAACTTGTCTGACACGCTCTCCGCGGCTCCGACAGAAATTAATTTCTGTCGGCTGC 1173
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
261 LeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGly 280
QY 1174 GGGCAACTGTTTTATTCCCGCCCGTTGCTCAGCCAATGGCGAGCCACCGTGAAGCTC 1233
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
281 GlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeu 300
QY 1234 TATACATGATCGAGAATGCTCAGCAGGATAGGGTGTGCTATCCCCACCATATCGAT 1293
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 TyrThrSerValGluAsnAlaGlnGlnAspLysGlyValAlaIleProHisAspIleAsp 320
QY 1294 CTTGTGATTTCGGGTGCTGCTATTTCAGGATTATGACACCCAGCATGACAGCATCGGCC 1353
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
321 LeuGlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGlnGlnAspArgPro 340
QY 1354 ACCCGCTCGCTCGCCCATCTCGGCCTTTTCTGTTCTCCGAGCAAAATGATGACTTTGG 1413
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
341 ThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyr 360
QY 1414 CTGTCCCTCACTGACCGCAGTATGACCACTCCACTTACGGGTGCTCACTGCGCGGTT 1473
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 LeuSerLeuThrAlaAlaGlnTyrAspGlnSerThrTyrGlySerSerThrGlyProVal 380
QY 1474 TATATCTCGACAGCGCTGACTTTGTTGATGTTGCGACTGGCGCGCAGCGCCGTAGCCCGA 1533
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
381 TyrIleSerAspSerValThrLeuValAsnValAlaThrGlyValGlnAlaValAlaArg 400
QY 1534 TCGCTTGAGTGTCCAAAGTCACTCCGAGCGGCGCCCTCCCGACTCTTGAGCAATAT 1593

```
|||||
Db 401 SerLeuAspTrpSerLysValThrLeuAspGlyArgProLeuProThrValGluGlnTyr 420
QY 1594 TCAAGACATCTTTGTCCTCCCTTCGGTGGCAAGCTCTCCCTTTGGGAGCCGGCACA 1653
Db 421 SerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThr 440
QY 1654 ACAAAAGCAGTTATCTTATAATATAATACTACTGTAGTACACAGATTCTGATGAA 1713
Db 441 ThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeuIleGlu 460
QY 1714 AATGCTCCCGCCATCGGGTGGCCATTTCAACCTATATACACAGGCTTGGGGCGGTCG 1773
Db 461 AsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrArgLeuGlyAlaGlyPro 480
QY 1774 GTCCGCATTTCTGGGGCGGGTTTTGGTCCACGCTCCGGCCCTGGCTCTGCTGAGGAT 1833
Db 481 ValAlaIleSerAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuLeuGluAsp 500
QY 1834 ACTTTTGATTATCGGGGCGGGCGCACACATTTGATGACTTCTGCCCTGAATGCCGCGCT 1893
Db 501 ThrPheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysArgAla 520
QY 1894 TTAGGCTCCAGGGTTGTCTTTCCAGTCAACTGTGCTGAGCTCCAGCGCTTTAAAGTT 1953
Db 521 LeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysVal 540
QY 1954 AAGTGGGTAAACTCGGGAGTTG 1977
Db 541 LysValGlyLysThrArgGluLeu 548
```

RESULT 8

```
US-08-542-634-14
; Sequence 14, Application US/08542634
; Patent No. 6214970
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; APPLICANT: McAtee, C. Patrick
; APPLICANT: Yarbrough, Patrice O.
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/542,634
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
```

```
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico Strain)
; INDIVIDUAL ISOLATE: ORF-2
US-08-542-634-14
Alignment Scores:
Pred. No.: 0 Length: 660
Score: 512.00 Matches: 658
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 1
Query Match: 77.69% Indels: 3
DB: 3 Gaps: 1
```

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-08-542-634-14 (1-660)

```
QY 1 ATCGCCCTAGGCTCTTTTGTCTGTTCTCTCTTGTCTTGTCTGTGCTGTGCTGCGCCGCA 60
Db 1 MetArgProArgProLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY 61 CCGACCCGTCAGCCGTCGTGGCGCGGCAGCGCGGTACCGCGGTGGT 120
Db 21 ProThrGlyGlnProSerGlyArgArgGlyArgSerGlyGlyThrGlyGlyGly 40
QY 121 TTCTGGGTGACCGGGTTGATTCACGCTTCCGAATCCCTATATATCATCAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTTGCCCGACGCTTGGCTGCGGTGCTGGAGCTGCGCTTCCGACACCGACGC 240
Db 61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY 241 CGGCACCTTGGCTCCTTGGCGAGATCAGGCCCGCGCCCTCCGCTGCCCTCCGCTCGC 300
Db 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
QY 301 CGACCTGCCACAGCGGGTGGCGGCTGACGGCTGTGGCGCTGCCCATGACACCTCA 360
Db 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCCGTCGGGACGTTGATTCGCGGTGCAATTCAGCCGCCAGTATATTTCTACT 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
QY 421 TCACCCCTGCATCCTCTGTGGCTCTGGCACTAATTTAGTCCTGTATGACGCCCTT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AATCCGCTCTGCGCTGCAGGACGGTACTAATACTCACATTATGGCCACAGAGCCTCC 540
Db 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCACAGTACCGGTTGCCGGCTACTATCCGTTACCGGCCCTAGTGCCTAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTCTGGGCTCAAAACACACACCCCTACA 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCTGTTGACATGAATTCATTACTTCCACTGATGTCAGGATCTGTGTTCAACCTGGCATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCATCTGAATTTGTCATCCGAGGAGGCTTCACTACCGCAATCAGGTTGGGCTCG 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTTGAGACATCTGCTGTTGCTGAGGAGGAAGCAACCTCCGCTCTGTGTCATGTTATGATA 840
Db 261 ValGlnThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCTCCAGTTAACTCCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 900
Db 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
```

```
QY 901 GACTTTGCTTGTAGAGCTTGTGAGTTTCGAAATCTCACCACTGTAAACACCAATACACGTGTG 960
Db |||||||
QY 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
Db |||||||
QY 961 TCCGTTACTCCAGCACTCTCTCGTCA----CTCCGCCCGAGGGCGCGAGGACTCGGA 1016
Db |||||||
QY 321 SerArgTyrSerThrAlaArgHisArgLeuArg-ArgGlyAlaAspGlyThrAlaG1 340
Db |||||||
QY 1017 GCTGACCACAATCGACGACACAGGTTTCATGAAGATCTCCACTTTACCGGCTTAATGG 1076
Db |||||||
QY 340 uleuthrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnG1 360
Db |||||||
QY 1077 GGTAGTGAAGTCGGCGGGGATAGTCTCTAACATTACTTAACTTGTCTGACACGCTCTCT 1136
Db |||||||
QY 360 yValGlyGluValGlyArgGlyLeuAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLe 380
Db |||||||
QY 1137 CGCGGGCTCCGACAGAAATTAATTCGTGCGCTGGCGGCAACTGTTTTATTCGCGCC 1196
Db |||||||
QY 380 uGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPr 400
Db |||||||
QY 1197 GGTGTCTCAGCAATGGCGAGCAACCGTGAAGTCTATACATCAGTGGAGAACTCTCA 1256
Db |||||||
QY 400 ovalValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaG1 420
Db |||||||
QY 1257 GCAGGATAAGGTTGTCTATCCCCCAGCATATCGATCTTGGTGATTCTCGGTGTGTCAT 1316
Db |||||||
QY 420 nGlnAspLysGlyValAlaAlaIleProHisAspLeuLeuGlyAspSerArgValVal1 440
Db |||||||
QY 1317 TCAGGATATGACACACCATGAGCAGATCGGCGCCACCGTGGCTGCGCTGCGCCATCTCG 1376
Db |||||||
QY 440 eGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerAr 460
Db |||||||
QY 1377 GCCTTTTCTGTTCTCCGAGCAATGATGATCTTTGGCTGCTCCCTCACTGCAGCGCGAT 1436
Db |||||||
QY 460 gProPheSerValLeuArgAlaAsnAspValLeuTyrPheSerLeuThrAlaAlaGluTyr 480
Db |||||||
QY 1437 TGACCAAGTCCACTTACGGGTCTGCTCACTGGCGCGGTTTATATCTCGACAGCGTGACTTT 1496
Db |||||||
QY 480 rAspGlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLe 500
Db |||||||
QY 1497 GGTGAATGTGGACGTGGCGCGAGCCGCTAGCCGATCGCTTGACTGTGTCGCAAGTCAAC 1556
Db |||||||
QY 500 uValAsnValAlaThrGlyAlaGlnAlaValaArgSerLeuAspTyrPheSerLysValTh 520
Db |||||||
QY 1557 CCTCGACGGCGGCCCTCCGACTGTTGAGCAATATCCAAAGACATCTTTGTGCTCC 1616
Db |||||||
QY 520 rLeuAspGlyArgProLeuProThrValGluGlnTyrSerLysThrPheValLeuPr 540
Db |||||||
QY 1617 CTTCTGTCGAAGCTCTCCTTTTGGAGCGCGGCACAAAGCAGGTTATCCTTATAA 1676
Db |||||||
QY 540 oLeuArgGlyLysLeuSerPheTyrPduAlaGlyThrThrLysAlaGlyTyrProTyrAs 560
Db |||||||
QY 1677 TTATAATACTACTGTAGTGACCAATCTGATTCGAAATGCTGCGCGCCATCGGTCGC 1736
Db |||||||
QY 560 nTyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAl 580
Db |||||||
QY 1737 CATTTCAACTATACACAGGCTTGGCGCGGTCGCGTTCGCAATTTCTGCGCGCGCGGT 1796
Db |||||||
QY 580 alIeSerThrTyrThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaAlaVa 600
Db |||||||
QY 1797 TTTGCTCCACGCTCCGCGCTGCTCTGCTGGAGGATACTTTGTATTATCCGGGCGGGC 1856
Db |||||||
QY 600 lLeuAlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAl 620
Db |||||||
QY 1857 GCACATTTGATGATCTCGCTGATGATCGCGCTTAGCCCTTCAGGTTTGCTTT 1916
Db |||||||
QY 620 aHisThrPheAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPh 640
Db |||||||
QY 1917 CCAGTCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTGCTTAAAGTCTCGGAGTT 1976
Db |||||||
QY 640 eGlnSerThrValAlaGluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLe 660
Db |||||||
```

```
QY 1977 G 1977
Db 660 u 660
RESULT 9
PCT-US95-13703-14
; Sequence 14, Application PC/TUS9513703
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
; TITLE OF INVENTION: USBS THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13703
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico Strain)
; INDIVIDUAL ISOLATE: ORF-2
PCT-US95-13703-14
Alignment Scores:
Pred. No.: 0 Length: 660
Score: 512.00 Matches: 658
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 1
Query Match: 77.69% Indels: 3
Db: 5 Gaps: 1
US-09-851-410A-10_COPY_5117_7096 (1-1980) x PCT-US95-13703-14 (1-660)
QY 1 ATGCGCCCTAGCCCTCTTTGCTGTGTTCTCTCTTTGTTCTGCTATGTTGCCCGGCCA 60
Db |||||||
QY 1 MetArgProArgProLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
Db |||||||
QY 61 CCGACCGGTACGCCGCTCTGCGCGCGCGTACCGCGCGGTACCGCGCGGTGCT 120
Db |||||||
QY 21 ProThrGlyGlnProSerGlyArgArgGlyArgSerGlyThrGlyGlyGly 40
Db |||||||
QY 121 TTTGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCCATATTCATCCCAACCAAC 180
Db |||||||
QY 41 PheTyrGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
Db |||||||
QY 181 CCCTTTGCCAGAGTTGCGGTCGCTCGCTTCGACCTCGCTTCGCCCAACCGACGC 240
Db |||||||
QY 61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
Db |||||||
```

QY 241 CGGCCACTTGCTCACTTGCGAGATCAGGCCCGCCCTCGCTGCTCCCTCGC 300
Db |||||
81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
QY 301 CGACCTGCCACAGCGGGGCTGCGCGCTGACGGCTGTGGCGCTGCCCATGACACCTCA 360
Db |||||
101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCCGTCGGAGCTGATTCTCGGGGTGCAATTCAGCCGCCAGTATAAATTGCTACT 420
Db |||||
121 ProValProAspValAspSerArgGlyAlaAlaLeuArgGlnTrpAsnLeuSerThr 140
QY 421 TCACCCCTGACATCTCTGTGGCTCTGCGCTCAATTTAGTCCTGTATGACGCCGCCCT 480
Db |||||
141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuThrAlaAlaProLeu 160
QY 481 AATCCGCTCTGCGCTGCGAGCGGTACTAATCTACATATTGGCCACAGAGCCCTCC 540
Db |||||
161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCACAGTACCGGTTGCCGCGCTACTATCCGTTACCGGCCCTAGTGCCTAAT 600
Db |||||
181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCAGTTCGAGGCTATGCTATATCCATTTCTTGGCTCAAAACACACACCCCTACA 660
Db |||||
201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCTGTTGACATGAATTCATTAATCTTCCACTGATGTCAGGATTCCTGTTCAACCTGGCATA 720
Db |||||
221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCATCTGAATGGTCATCCCAAGGAGCGCTTCACTACCGCAATCAAGTTGGCGCTCG 780
Db |||||
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTTCAGACATCTGTGTGTGTGAGGAGAGCCACTCCGCTCTGTGATGTTATGCATA 840
Db |||||
261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCTCCAGTTAACTCTCATACCAATACCCCTTATACCGTGCCTTTGGCTTACTG 900
Db |||||
281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCTTAGAGCTTGAGTTCCGCAATCTCACCTGTACCACTGTACACCAATACAGTGTG 960
Db |||||
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTTACTCCAGCACTGCTCGTCA---CTCCGCGCGAGGGCGGACGGACTCGCGA 1016
Db |||||
321 SerArgTyrSerSerThrAlaArgHisArgLeuArg-ArgGlyAlaAspGlyThrAlaGln 340
QY 1017 GCTGACCACAACTGACGCCAGCGTTTCATGAAAGATCTCCACTTTACCGGCCTTAATGG 1076
Db |||||
340 uLeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGln 360
QY 1077 GGTAGGTGAAGTCCGCGCGGAGTAGCTTAACATTACTTAACCTTGCTGACAGCTCCT 1136
Db |||||
360 yValGlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLe 380
QY 1137 CGGCGGGCTCCCGACAGAAATTAATTTCTCGCTGCGCGGCAACTGTTTATTCGCGGCC 1196
Db |||||
380 uGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlnLeuPheTyrSerArgPr 400
QY 1197 GGTGTCTTCAGCCAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCA 1256
Db |||||
400 oValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1257 GCAGATGAAGGTGTGTATCCCGCCACAGATATCGATCTTGGTGTATTCGGTGTGGTCAAT 1316
Db |||||
420 nGlnAspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValValIle 440
QY 1317 TCAGGATTATGACACCAAGCATGACAGGATCGGCCCAACCGCTCGCTCGCCATCTCG 1376

Db 440 eGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1377 GCCTTTTCTGTTCTCCGAGCAATGATGACTTTTGGCTGTCCCTCACTGACGCGGAGTA 1436
Db |||||
460 gProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1437 TGACCACTGCACATTACGGGTGCTCAACTGGCCCGGTTTATATCTCGGACAGCGTGACTTT 1496
Db |||||
480 rAspGlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLe 500
QY 1497 GGTGAATGTTGCGACTGCGCGCAGGCGGTAGCCCGATCGCTTGACTGGTCAAAAGTCAC 1556
Db |||||
500 uValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValTh 520
QY 1557 CCTCGACGGGGCGGCCCTCCCGACTGTTGAGCAATATTCGAAGACATCTTTGTGCTCCC 1616
Db |||||
520 rLeuAspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuPr 540
QY 1617 CTTTCGTGGCAAGCTCTCTTTTGGGAGCGCGCACAAACAAAGCAGGTTATCCTTATAA 1676
Db |||||
540 oLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAS 560
QY 1677 TTATAATACTACTGCTAGTCACCAAGATTCTGATTGAAAATGCTCCGCGCCATCGGGTCGC 1736
Db |||||
560 nTyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAl 580
QY 1737 CATTTCAACCTATACACACAGGCTTGGGCGCGCTCGCTCCCATTTCTGCGCGCGCGGT 1796
Db |||||
580 aIleSerThrTyrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaAlaVa 600
QY 1797 TTTTCGCTCCAGCTCCGCTCCGCTCTGCTGAGAGTACTTTTGATTTATCCGGGCGGGC 1856
Db |||||
600 lLeuAlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAl 620
QY 1857 GCACACATTTTGATCACTTCTGCCCTGAATGCGCGCTTTAGGCTCCAGGGTGTGTGCTTT 1916
Db |||||
620 aHisThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPh 640
QY 1917 CCAGTCAACTGCTCGCTGAGTCCAGCGCTTAAAGTTAAAGTGGGTAAACCTCGGGAGTT 1976
Db |||||
640 eGlnSerThrValAlaGluLeuGlnArgLeuLysValGlyLysThrArgGluLe 660
QY 1977 G 1977
Db 660 u 660
RESULT 10
US-09-462-606-48
; Sequence 48, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
; FILE REFERENCE: 20264267U1
; CURRENT APPLICATION NUMBER: US/09/462,606
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-48
Alignment Scores: 0 Length: 660
Pred. No.:

APPLICANT: Bradley, Daniel W.
APPLICANT: Two, Jr-Shin
APPLICANT: Purdy, Michael A.
APPLICANT: Tam, Albert W.
APPLICANT: Krawczynski, Krzysztof Z.
APPLICANT: Yarbough, Patrice D.
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/259,148A
FILING DATE: 13-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APR-1983
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: C2, MEXICO, FIGURE 9
US-08-259-148A-18

Alignment Scores:
Pred. No.: 0 Length: 435
Score: 435.00 Matches: 435
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.01% Indels: 0
DBs: 1 Gaps: 0

US-09-851-410a-10_COPY_5117_7096 (1-1980) x US-08-259-148A-18 (1-435)

QY 673 AATTCATTACTTCACGTGATCTCAGGATCTTGTTCAACCTGGCATAGCACTGAATTG 732
DB 1 AsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGluLeu 20

QY 733 GTCATCCCAAGCGAGCGCCTTCACTACGCAATCAAGGTTCGGCGCTCGGTTGAGACATCT 792
DB 21 ValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThrSer 40
QY 793 GGTGTTGCTGAGGAGGAAGCCACCTCCGGTCTTGTTCATGTTATGCATACATGGCTCTCCA 852
DB 41 GlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySerPro 60
QY 853 GTTAACCTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGGACTTTCGCTTA 912
DB 61 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 80
QY 913 GAGCTTGAGTTTCGCAATCTCACACCTGTAAACCAATACACAGTGTCTCCGGTTACTCC 972
DB 81 GluLeuGluPheArgAsnLeuThrCysAsnThrAsnThrArgValSerArgTyrSer 100
QY 973 AGCACTGCTGCTCACTCCGCCGAGGGCCGACGGAGCTCGGAGCTCACCACCACTCCA 1032
DB 101 SerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeuThrThrAla 120
QY 1033 GCCACCAAGTTCATGAAGATCTCCACTTTACCGGCCCTTAATGGGTAGGTGAAGTCGGC 1092
DB 121 AlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyValGlyGluValGly 140
QY 1093 CGCGGATAGCTCTAAACATTACTTAACTTCTGACACGCTCTCCGGGGGCTCCCGACA 1152
DB 141 ArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThr 160
QY 1153 GAATTAATTTGCTGCGGTGGCGGCACTGTTTTATTCGCCCGGTTGTCTCAGCCCAAT 1212
DB 161 GluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsn 180
QY 1213 GCGAGCCCAACGCTGAAGCTCTATACATCAGTGGAGAATGCTCAGCAGATAGGGTGT 1272
DB 181 GlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGlnAspLysGlyVal 200
QY 1273 GCTATCCCCCACGATATCGATCTTGGTGATTTCGGTGGTGGTTCATTACAGATTATCACAAC 1332
DB 201 AlaIleProHisAspIleAspLeuGlyAspSerArgValIleGlnAspTyrAspAsn 220
QY 1333 CAGCATGAGCAGGATCGGCCCAACCCCGCTCGCCATCTCGGCTTTTCTGTTCTC 1392
DB 221 GlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeu 240
QY 1393 CGAGCAATGATGATCTTTGCTGCTCTCCTCAGCGCGGATGATGACAGTCCACTTAC 1452
DB 241 ArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyr 260
QY 1453 GGGTCGTCAACTGGCCCGGTTTATATCTCGACACGCTGACTTTGGTGAATGTTCCGACT 1512
DB 261 GlySerSerThrGlyProValTyrIleSerAspSerValThrLeuValAsnValAlaThr 280
QY 1513 GCGCGCAGGCGGTAGCCGATCGCTTCAGCTGGTCCAAAGTCACCTCGAGGGCGGCC 1572
DB 281 GlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeuAspGlyArgPro 300
QY 1573 CTCGGACTGTTGACCAATATTCAGACATTCTTTGCTCCCTCCCTTCGTTGGGAGCTC 1632
DB 301 LeuProThrValGluGlnTyrSerLysThrPhePheValLeuProLeuArgGlyLysLeu 320
QY 1633 TCCTTTTGGAGGCGGCACACAAAGCAGGTTATCTTATAATATAATACTACTGCT 1692
DB 321 SerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAla 340
QY 1693 AGTGACCAAGATTCTGATTGAAAAATGCTCGCGGCCATCGGGTCGCCATTTCACACCTATACC 1752
DB 341 SerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThr 360
QY 1753 ACCAGGCTTGGGCGCGTCCGGTCCGCAATTTCTGGCGCGCGGTTTGGCTCCAGCTCC 1812
DB 361 ThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeuAlaProArgSer 380

```
QY 1813 GCGCTGCTCTGCTGAGGATACCTTTTGATTATCCGGGGCGGCGCACATTTGATGAC 1872
Db 381 AlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHisThrPheAspAsp 400
QY 1873 TTCTGCCCTGAATGCCGCTTTAGCCCTCCAGGTTGTGCTTTCCAGTCAACTGTCGCT 1932
Db 401 PheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAla 420
QY 1933 GAGCTCCAGCGCTTAAGCTTAAGTGGGTAAACTCCGGAGTTG 1977
Db 421 GluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 435
RESULT 12
US-08-484-054-18
; Sequence 18, Application US/08484054
; Patent No. 5770689
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,054
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 870,985
; FILING DATE: 20-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: C2, MEXICO, FIGURE 9
US-08-484-054-18
Alignment Scores:
Pred. No.: 0 Length: 435
Score: 435.00 Matches: 435
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.01% Indels: 0
DB: 1 Gaps: 0
US-09-851-410a-10_COPY_5117_7096 (1-1980) x US-08-484-054-18 (1-435)
QY 673 AATTCCATTACTTCCACTGATGTCCAGATTCTTGTTCACACCTGGCATAGCATCTGAATTG 732
Db 1 AsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGluLeu 20
QY 733 GTCAATCCCAAGCGAGCGCTTCACTACCCCAATCAAGTTGGCGCTCGGTGAGACATCT 792
Db 21 ValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThrSer 40
QY 793 GGTGTTGCTGAGGAGGAAGCCACCTCCGCTCTGTTCATGTATGCATACATACATGCTCTCCA 852
Db 41 GlyValAlaGluGluGluAlaThrSerGlyLeuValMetCysIleHisGlySerPro 60
QY 853 GTTAACTCCTATACCAATACCCCTTATACCGGTGGCCCTTGGCTTACTGGACTTTGCCTTA 912
Db 61 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 80
QY 913 GAGCTTGAGTTTCGCAATCTCACCACCTGTAAACCAATACACAGTGTGTCCGTACTCC 972
Db 81 GluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgValSerArgTyrSer 100
QY 973 AGCACTGCTGCTCACTCCGCCCGAGGGCGGAGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGC 1032
Db 101 SerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeuThrThrThrAla 120
QY 1033 GCCACCGAGTTTCATGAAAGATCTCCACTTTACCGGCTTTAAATGGGTAGTGAAGTCGCGC 1092
Db 121 AlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyValGlyGluValGly 140
QY 1093 CCGGGGATAGCTTAACATTACTTAACCTTGTGTGACACGCTCTCTCGGGGGCTCCGGACA 1152
Db 141 ArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThr 160
QY 1153 GAAATTAATTTGCTCGCTGGCGGCAACTGTTTTTATTCGCCGCCGCTTCTCTCAGCAAT 1212
Db 161 GluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsn 180
QY 1213 GCGAGCAACCGTGAAGCTCTATACATCAGTGGAGAATGCTCAGCAGGATAAGGTGTT 1272
Db 181 GlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnAlnAspLysGlyVal 200
QY 1273 GCTATCCCCCAGCATATCGATCTTGGTGTATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1332
Db 201 AlaIleProHisAspIleAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsn 220
QY 1333 CAGCATGAGCAGGATCGGCCACCCCGCTCGCTCGCCCATCTCGGCTTTTCTGTCTTC 1392
Db 221 GlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeu 240
QY 1393 CGAGCAATGATGTACTTTGGCTGCTCCTCACTGCGAGCGAGTATGACAGTCCACTTAC 1452
Db 241 ArgAlaAsnAspValLeuTriLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyr 260
QY 1453 GGTGCTCAACTGGCCCGGTTTATATCTCGGACACGGTACTTTGTTGATGTTGCGACT 1512
Db 261 GlySerSerThrGlyProValTyrIleSerAspSerValThrLeuValAsnValAlaThr 280
```

1513 GCGCGCAGCGCGTAGCCGATCGCTTGATGCTCAAAAGTCACCCCTCGACGGCGCGCC 1572
Db GlyAlaGlnAlaValAlaArgSerLeuAspTTPSerLysValThrLeuAspGlyArgPro 300
1573 CTCGCCAGCTGTGAGCAATATCCAGACATCTTTGTGCTCCCTTCCTGTCGCAAGCTC 1632
Db LeuProThrValGluGlnTyrSerLysThrPheValLeuProLeuArgGlyLysLeu 320
1633 TCCTTTTGGGAGCGCGGCAACAAAGCAGGTTATCTCTTATAATTAATACTACTGCT 1692
Db SerPheTTPGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnLysThrThrAla 340
1693 ATGACCAAGATCTTGATGAAATGCTCGCGCATCGGTCGCCATTTCAACCTATACC 1752
Db SerAspGlnLeuLeuIleGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThr 360
1753 ACCAGGCTGGCGCGGTGCGGTCGCATTTCTGCGCGCGCGGTTTGTGCTCAGCTCC 1812
Db ThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeuAlaProArgSer 380
1813 GCGCTGGCTCTGCTGAGGATACTTTTGATTTATCGGGCGGCGGCACACATTTTGATGAC 1872
Db AlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHisThrPheAsp 400
1873 TTCTGCCCTGAATCGCGCTTTAGCCCTCCAGGTTGTGCTTCCAGTCAACTGTGCT 1932
Db PheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAla 420
1933 GAGCTCCAGCGCTTAAAGTTAAGTGGTAAACTCGGAGTTG 1977
Db GluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 435

RESULT 13

US-07-876-941A-18
; Sequence 18, Application US/07876941A
; Patent No. 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; TITLE OF INVENTION: Antibodies
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/876,941A
; FILING DATE: 01-MAY-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.33
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: C2, MEXICO, FIGURE 9
; US-07-876-941A-18
Alignment Scores:
Pred. No.: 0 Length: 435
Score: 435.00 Matches: 435
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.01% Indels: 0
DB: 2 Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-07-876-941A-18 (1-435)
QY 673 AATTCATTACTTCCATGATGTCAGGATTCCTGTTCAACCTGGCATAGCATCTGAATG 732
Db 1 AsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGluLeu 20
QY 733 GTCATCCCAAGCGAGCGGCTTCACCTACCGCAATCAAGGTTGGCGCTCGGTGAGCATCT 792
Db 21 ValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThrSer 40
QY 793 GGTGTGCTGAGGAGGAGGAGCCACCTCCGCTCTTGTTCATGTTATGATACATCGCTCTCCA 852
Db 41 GlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySerPro 60
QY 853 GTTAACTCCTATACCAATACCCCTTATACCGGTGCCCTTACTGGACTTTGCTTTA 912
Db 61 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 80
QY 913 GAGCTTGAGTTTCGCAATCTCACCACCTGTAACCAATACACGTGTGTCCTGTTACTCC 972
Db 81 GluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgValSerArgTyrSer 100
QY 973 AGCACTGCTGCTCACTCCGCGCGGAGGCGGAGCTCGGAGCTGACCACTGCA 1032
Db 101 SerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeuThrThrAla 120
QY 1033 GCCACCAAGTTTCATGAAAGATCTCCACTTTACCGGCTTAATGGGGTAGGTGAAGTCGGC 1092
Db 121 AlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyValGlyGluValGly 140
QY 1093 CGCGGATAGCTTAACATTTACTTTAACTTCTGACACGCTCTCGCGGGGCTCCGACA 1152
Db 141 ArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThr 160
QY 1153 GAATTAATTTTCGTGCGTGGCGGCAACTGTTTATTCCCGCCCGTGTGCTCAGCCAAT 1212
Db 161 GluLeuSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsn 180
QY 1213 GCGAGCAACCGTGAAGCTCTTATACATCAGTGGAGAAATGCTCAGCAGGATAAGGGTGT 1272
Db 181 GlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGlnAspLysGlyVal 200

QY	1273	GCTATCCCCAGCATATCGATTCTTGCTGATTGCGGTGGTCAATTCAGGATTATGACAAC	1332	;	CURRENT APPLICATION DATA:
Db	201	AlaileProHisAspIleAspIleAspGlyAspSerArgValValIleGlnAspIleAspAsm	220	;	APPLICATION NUMBER: US/07/870,985A
QY	1333	CAGCATGAGCAGATCGGCCACCCCGTCCCTCGCCCATCTCGGCTTTTCTGTCTC	1392	;	FILING DATE: 20-APRIL-1992
Db	221	GlnHisGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeu	240	;	CLASSIFICATION: 435
QY	1393	CGAGCAAAATGATATCTTTGGCTGTGCTCCTCACTGCGCGGAGATGACGAGTCCACTTAC	1452	;	PRIOR APPLICATION DATA:
Db	241	ArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTrpAspGlnSerThrTyr	260	;	APPLICATION NUMBER: US 822,335
QY	1453	GGGTGCTCAACTCGCGCGTTTATCTCGAGAGCTGACTTTGGTGAATGTCGACT	1512	;	FILING DATE: 17-JAN-1992
Db	261	GlySerSerThrGlyProValIleSerAspSerValThrLeuValAsnValAlaThr	280	;	PRIOR APPLICATION DATA:
QY	1513	GGCGCGCAGCCCTAGCCCATCGCTTGACTGCTGCCAAAGTCACCCCTCGAGCGCGGCC	1572	;	APPLICATION NUMBER: US 505,888
Db	281	GlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeuAspGlyArgPro	300	;	FILING DATE: 05-APRIL-1990
QY	1573	CTCCGACTGTTGAGCAATATCCAGACATCTCTTGTGCTCCGCTTCTGTCGCAAGCTC	1632	;	PRIOR APPLICATION DATA:
Db	301	LeuProThrValGluGlnTrpSerLysThrPhePheValLeuProLeuArgGlyLysLeu	320	;	APPLICATION NUMBER: US 420,921
QY	1633	TCCTTTTGGAGCGCGCACAAAGCAGGTTATCTTATATATATATATATATCTACTGCT	1692	;	FILING DATE: 13-OCTOBER-1989
Db	321	SerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProLysAsnThrThrAla	340	;	PRIOR APPLICATION DATA:
QY	1693	AGTGACCAATCTGATTGAAATGCTGCGCGCATCGGTCGCCATTTCAACCTATACC	1752	;	APPLICATION NUMBER: US 367,486
Db	341	SerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThr	360	;	FILING DATE: 16-JUNE-1989
QY	1753	ACGAGCTTGGCGCGTCCGCTCGCTTCTGCGCGCGCTTTGGCTCCACGCTCC	1812	;	PRIOR APPLICATION DATA:
Db	361	ThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeuAlaProArgSer	380	;	APPLICATION NUMBER: US 336,672
QY	1813	GCCTGCTCTGCTGGAGGACTTTTGATTATCCGGCGCGCGGACACATTTGATGAC	1872	;	FILING DATE: 11-APRIL-1989
Db	381	AlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHisThrPheAspAsp	400	;	PRIOR APPLICATION DATA:
QY	1873	TTCTGCCCTGAATGCCCGCTTTAGCGCTCCAGGTTGTGCTTCCAGTCACTGCTGCT	1932	;	APPLICATION NUMBER: US 208,997
Db	401	PheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAla	420	;	FILING DATE: 17-JUNE-1988
QY	1933	GAGCTCAGCGCTTAAGTTAAGTGGGTAAACTCGGAGGTG 1977		;	ATTORNEY/AGENT INFORMATION:
Db	421	GluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 435		;	NAME: Fabian, Gary R.
RESULT 14					REGISTRATION NUMBER: 33,875
US-07-870-985A-18					REFERENCE/DOCKET NUMBER: 4600-0093.30
Sequence 18, Application US/07870985A					TELECOMMUNICATION INFORMATION:
Patent No. 6455492					TELEPHONE: (415) 324-0880
GENERAL INFORMATION:					TELEFAX: (415) 324-0960
APPLICANT: Reyes, Gregory R.					SEQUENCE CHARACTERISTICS:
APPLICANT: Bradley, Daniel W.					LENGTH: 435 amino acids
APPLICANT: Twu, Jr-Shin					TYPE: amino acid
APPLICANT: Purdy, Michael A.					TOPOLOGY: linear
APPLICANT: Tam, Albert W.					MOLECULE TYPE: protein
APPLICANT: Krawczynski, Krzysztof Z.					HYPOTHETICAL: NO
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method					ORIGINAL SOURCE:
NUMBER OF SEQUENCES: 22					INDIVIDUAL ISOLATE: C2, MEXICO, FIGURE 9
CORRESPONDENCE ADDRESS:					US-07-870-985A-18
ADDRESSEE: Dehlinger & Associates					Alignment Scores:
STREET: 350 Cambridge Avenue, Suite 250					Pred. No.: 0
CITY: Palo Alto					Score: 435.00
STATE: CA					Matches: 435
COUNTRY: USA					Conservative: 0
ZIP: 94306					Best Local Similarity: 100.00%
COMPUTER READABLE FORM:					Query Match: 66.01%
MEDIUM TYPE: Floppy disk					DB: 4
COMPUTER: IBM PC compatible					
OPERATING SYSTEM: PC-DOS/MS-DOS					
SOFTWARE: PatentIn Release #1.0, Version #1.25					

QY	1033	GCCACCAAGGTTTCATGAAAGATCTCCACTTTACCGGCTTTAAATGGGGTAGGTGAAGTCGGC	1092
DB	121	AlaThrArgPheMetIysAspLeuHisPheThrGlyLeuAsnGlyValGlyGluValGly	140
QY	1093	CGGGGATAGCTCTAAACATTACTTAACCTTGTGACACAGCTCTCCGGCGGGTCCCGACA	1152
DB	141	ArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThr	160
QY	1153	GAATTAATTTTCGTGCGTGGCGGCAACTGTTTTATTCCGCCCGGTCGTCACAGCAAT	1212
DB	161	GluLeuIleSerSerAlaGlyGlnLeuPheTyrSerArgProValValSerAlaAsn	180
QY	1213	GGCGAGCAACCGTGAAGCTCTATACATCAGTCGAGAACTCTCAGCAGGATAAGGCGTGT	1272
DB	181	GlyGluProThrValIysLeuTyrThrSerValGluAsnAlaGlnAspIysGlyVal	200
QY	1273	GCTATCCCCACGATATCGATCTTTGGTGATTCCGCTGTGTCATTCAGGATTATGACAAC	1332
DB	201	AlaIleProHisAspIleAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsn	220
QY	1333	CAGCATGAGCAGGATCGGCCACCCCCGTGCGCTCGCCATCTCGGCTTTTTCTGTCTCC	1392
DB	221	GlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeu	240
QY	1393	CGAGCAATGATGTACTTTGGCTGCCTCCTCACTCAGCCGAGTATGACCACTCCACTTAC	1452
DB	241	ArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyr	260
QY	1453	GGGTGCTCAACTGGCGCCGGTTTTATCTTCGGACAGCGTGACTTTGGTGAATGTGCGACT	1512
DB	261	GlySerSerThrGlyProValTyrIleSerAspSerValThrLeuValAsnValAlaThr	280
QY	1513	GGCGCGCAGGCGGTAGCCCGATCGCTTGACTGTGTCCAAAGTCACCTTCGAGGGCGGCC	1572
DB	281	GlyAlaGlnAlaValAlaArgSerLeuAspTrpSerIysValThrLeuAspGlyArgPro	300
QY	1573	CTCCGACCTGTTGACCAATATCCAGACATTCTTTGTGCTCCCCCTTCGTGGCAAGCTC	1632
DB	301	LeuProThrValGluGlnTyrSerIysThrPhePheValLeuProLeuArgGlyIysLeu	320
QY	1633	TCCTTTTCGGAGCGCGGCACAAACACAGGTTATCCTTATAATTATACTACTGCT	1692
DB	321	SerPheTrpGluAlaGlyThrThrIysAlaGlyTyrProTyrAsnTyrAsnThrThrAla	340
QY	1693	AGTGACCAAGATTCTGATTGAAATGCTCGCGCCCATCGGGTCGCCATTTCAACCTATACC	1752
DB	341	SerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThr	360
QY	1753	ACCAGGTTGGGGCGGTTCGGTCCGCATTTCTGGCGCGCGGTTTGGCTCCACGCTCC	1812
DB	361	ThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeuAlaProArgSer	380
QY	1813	GCCTGGCTCTGCTCGAGGATCTTTGATTATCCGGCGCGCGGCACACATTTTGATGAC	1872
DB	381	AlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHisThrPheAspAsp	400
QY	1873	TTCTGCGCTTGAATCGCGCGCTTTAGCGCTTCCAGGTTGTGCTTTCCAGTCAACTGTGCT	1932
DB	401	PheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAla	420
QY	1933	AGACTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTTG	1977
DB	421	GluLeuGlnArgLeuIysValIysValGlyIysThrArgGluLeu	435

```

1  APPLICANT: Zhang, Yifan
2  TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
3  NUMBER OF SEQUENCES: 31
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE: Dehlinger & Associates
6  STREET: 350 Cambridge Ave., Suite 250
7  CITY: Palo Alto
8  STATE: CA
9  COUNTRY: USA
10 ZIP: 94306
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: PatentIn Release #1.0, Version #1.25
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/08/542,634
19 FILING DATE:
20 CLASSIFICATION: 435
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Fabian, Gary R.
23 REGISTRATION NUMBER: 33,875
24 REFERENCE/DOCKET NUMBER: 4600-0293.30
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: (415) 324-0880
27 TELEFAX: (415) 324-0960
28 INFORMATION FOR SEQ ID NO: 16:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH: 549 amino acids
31 TYPE: amino acid
32 STRANDEDNESS: single
33 TOPOLOGY: linear
34 MOLECULE TYPE: protein
35 HYPOTHETICAL: NO
36 ORIGINAL SOURCE:
37 INDIVIDUAL ISOLATE: Hepatitis E virus (Mexico strain)
38 INDIVIDUAL ISOLATE: r62kDa, FIGURE 4
39 US-08-542-634-16
40
41 Alignment Scores:
42 Pred. No.: 0 Length: 549
43 Score: 401.00 Matches: 547
44 Percent Similarity: 99.45% Conservative: 0
45 Best Local Similarity: 99.45% Mismatches: 1
46 Query Match: 60.85% Indels: 3
47 DB: 3 Gaps: 1

```

```
QY 694 GTGAGGATCTTGTTCACCTGGCATAGCATCTGAATTGGTCATCCCAAGCGAGCGGCTT 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTACCGCAATCAAGGTGGCGTTCGAGTTCAGACATCTGGTGTGCTGAGGAGAACCC 813
Db 141 HisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCCGGTCTGTGCATGTTATGTCATATGCTCTCCAGTTAACTCCTATACCAATACC 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThr 180
QY 874 CCTTATACCGGTGCGCTTACTGACTTTCCTTAGAGCTTGGTGTGCTTTCGCAATCTC 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuPheArgAsnLeu 200
QY 934 ACCACTGTAAACACCAATACACGTGTCTCCGCTTACTCCAGCATCTCGTCA----CTC 989
Db 201 ThrThrCysAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
QY 990 CGCCCGAGGGCGGCGGACTCGGAGCTGCACCACTGCAGCCACCAAGGTTCAATGAA 1049
Db 221 Arg-ArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLeu 240
QY 1050 AGATCTCCACTTTACCGGCTTAATGGGTAGGTAGTGGCGCGGAGTACTCTAAC 1109
Db 240 sAspLeuHisPheThrGlyLeuAsnGlyValGlyGluValGlyArgGlyIleAlaLeuTh 260
QY 1110 ATTACTTAACCTTGTCTGACACGCTCCTCGCGCGGCTCCCGACAGAAATTAATTCGTCGC 1169
Db 260 rLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAl 280
QY 1170 TGGCGGGCACTGTTTATTCGCGCGGTTGCTCAGCCAATGGCGAGCCAAACCGTGAA 1229
Db 280 aGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrVally 300
QY 1230 GCTCTATACATCAGTGGAGATCTCAGCAGGATAGGGTGTGCTATCCCCCACCAGTAT 1289
Db 300 sLeuTyrThrSerValGluAsnAlaGlnAspLysGlyValAlaIleProHisAspI 320
QY 1290 CGATCTTGTGTGTCGCTGTCATTGAGATATGACACAGCATGAGCAGGATCG 1349
Db 320 eAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspAr 340
QY 1350 GCCCACCCTGCTCGCCCATCTCGCGCTTTCTGTTCTCCGAGCAATGATGACT 1409
Db 340 gProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValle 360
QY 1410 TTGGTGTCCCTCCTCAGCGAGTATGACAGTCCACTTACGGTCTGCTCACTGGGCC 1469
Db 360 uTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPr 380
QY 1470 GGTTTATATCTCGGACAGCGTGACTTTGGTGAATGTTGCGACTGGCGCGCAGCGGTAGC 1529
Db 380 oValTyrIleSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAl 400
QY 1530 CCGATCGGTGACTGCTCAAGTCAACCTCAGCGGCGGCCCTCCCGACTGTTTGAGCA 1589
Db 400 aArgSerLeuAspTrpSerLysValThrLeuAspGlyArgProLeuProThrValGluGl 420
QY 1590 ATATTCCAAAGACATTTCTTGTGCTCCCTTGTGCGAGCTCTCTCTTTGGAGCGCG 1649
Db 420 nTyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGl 440
QY 1650 CACAACAAAGCAGGTTATCCCTATATTAATACTACTGCTAGTACCAGATTCTGAT 1709
Db 440 yThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeuIl 460
QY 1710 TGAAATGCTCGCGCCATCGGTCGCGCATTTCAACCTATACCAAGCTTGGGCGCG 1769
Db 460 eGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrThrArgLeuGlyAlaGl 480
QY 1770 TCCGCTCGCCCATTTCTGCGGCGCGGCTTTGGCTCCAGCTCCGCGCTGGCTCTGCTGA 1829
```

```
Db 480 yProValAlaIleSerAlaAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuGl 500
QY 1830 GGATACCTTTTCATTATCCGGGGCGGCGCACACATTTGATGACTTCTGCCCTGAATGCCG 1889
Db 500 uAspThrPheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysAr 520
QY 1890 CGCTTTAGGCTCCAGGTTGTGCTTTCAGTCACTGCTCAGCTCCAGCTCCAGCGCCTTAA 1949
Db 520 gAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLy 540
QY 1950 AGTTAAGGTGGTAAAACTCGGGAGTTG 1977
Db 540 sValLysValGlyLysThrArgGluLeu 549
```

Search completed: August 16, 2004, 14:00:12
Job time : 52.9849 secs